

DE GRUYTER

*Stephen R. Palmquist (Ed.)*

**CULTIVATING  
PERSONHOOD:  
KANT AND ASIAN  
PHILOSOPHY**

## Cultivating Personhood: Kant and Asian Philosophy



# Cultivating Personhood: Kant and Asian Philosophy

Edited by  
Stephen R. Palmquist

De Gruyter

ISBN 978-3-11-022623-2  
e-ISBN 978-3-11-022624-9

*Library of Congress Cataloging-in-Publication Data*

Cultivating personhood : Kant and Asian philosophy / edited by  
Stephen R. Palmquist.

p. cm.

Proceedings of a conference held in May 2009 in Hong Kong.

Includes bibliographical references and index.

ISBN 978-3-11-022623-2 (hardcover : alk. paper)

1. Persons – Congresses. 2. Philosophical anthropology – Congresses. 3. Kant, Immanuel, 1724–1804 – Congresses. 4. Philosophy, Asian – Congresses. I. Palmquist, Stephen R.

BD450.C85 2010

128–dc22

2010036203

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2010 Walter de Gruyter GmbH & Co. KG, Berlin/New York  
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen  
∞ Printed on acid-free paper

Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

To Philadora Grace

A genuine union of East and West



## Preface

Chinese philosophy first came to my attention in the early 1980s, not long after I read Kant's three *Critiques* in March of 1981. Even at that early stage, I felt instinctively that Kant is deeply Chinese in his thinking, with close parallels (in very different ways) to both Confucian and Daoist philosophers. Soon after coming to Hong Kong to teach in 1987, I learned of Mou Zongsan and his intriguing challenge both to his own Chinese tradition (to take Kant more seriously) and to Western Kant scholars (to make use of Chinese philosophical resources to fill gaps in philosophical wisdom left by Kant). I gradually became convinced that a major international conference was not only possible, but increasingly urgent as a path to forge in cross-cultural philosophy, given the rising place of China on the world scene. Belonging to a department uncommonly blessed with scholars well-versed in Kant's philosophy gave me confidence to believe that my own university would make an ideal venue for such a ground-breaking event. Yet it took over sixteen years to get past the idea stage.

Not until May of 2007, when Prof. Cheng Chung-ying visited the Department of Religion and Philosophy at Hong Kong Baptist University as a scholarly consultant and gave his resounding approval to the idea, did the planning process for the "Kant in Asia" conference actually begin. On a return visit later that year, Prof. Cheng suggested that the conference focus on the unity of personhood, a theme that was eventually adopted when the conference Organizing Committee held its first meeting in January of 2008. On the first day of the Chinese Year of the Rat, a preliminary Call for Papers was sent to Kant scholars around the world, and the response was so overwhelming that the Committee kept needing to increase the planned number of participants. The assistance of numerous philosophical societies around the world was essential in promoting the conference so effectively. In the end over 200 abstracts were submitted. With funding kindly promised by the Department in March of 2008 and by the University a few months later, we eventually invited three distinguished keynote speakers and 94 paper presenters, from over 30 different countries.

As the Year of the Ox drew near (early 2009), almost as if to insure that our ideals of unity could not be realized too fully, the Organizing



Committee began experiencing something like a Kantian “conflict of the faculties” first hand. Procedural tensions eventually led a minority on the Committee to stage a mutiny, casting a shadow over the Department’s involvement. Fortunately, the majority of my colleagues, together with a team of over 20 students, stuck with their commitments, providing much-needed assistance in the weeks preceding the long-awaited event. With the many challenges we faced together in the days leading up to the conference, the second of the two sentences quoted at the beginning of the Editor’s Introduction became the motto for the major international conference that took place from 20–23 May 2009. “Kant in Asia: The Unity of Human Personhood” was a great success by all counts, even regarded by some participants as instigating a sea change in global Kant studies.

Not long after the conference, Walter de Gruyter offered to publish the proceedings. The present volume includes revised versions of the three keynote lectures, followed by 64 out of approximately 80 contributed papers that were submitted after the conference. In editing these essays I have sought to strike a balance between unity of form and diversity of content. Certain stylistic and grammatical standards have been applied to the essays, and I would like to thank each contributor for her or his forbearance in putting up with the requirements of the chosen conventions, even when these conflicted with the author’s own preferred usages. Several features, however, proved so distinctive of different cultural approaches to the issues being discussed that I elected not to impose a common standard. Perhaps the best example is the use of Chinese (and other Asian) names. As the form of these names, when expressed in European languages, can vary widely, and as names are among the most personal of all words, I have allowed each author to name both him/herself and others in whatever manner she or he prefers. The same goes for names of philosophies or traditions (e.g., “Daoism” vs. “Taoism”), except that I have ensured that all such names are capitalized. A brief biographical sketch of each contributor can be found at the end of the book. Following this Preface is an explanation of the referencing system used throughout the book.

I would like to thank some of the many persons who, following Prof. Cheng’s crucial initial support, made this project possible. Without the backing of Kwan Kai Man, my Department Head in 2007–2008, the conference never would have materialized. The colleagues who joined me on the Organizing Committee (Jonathan, Leo, William, Kwok Kui, and Ellen) each made crucial contributions in shaping the

program as it developed over a period of more than a year; despite the conflict that arose, I will always cherish the collegial trust we six developed during those initial months of planning. For encouraging the University to provide a substantial additional grant, shortly after the conference first became a Department event, and for continuing their support by appearing as honored guests at the Opening Ceremony, I thank HKBU's Arts Faculty Dean, Chung Ling, and Vice President for Research and Institutional Development, Tsoi Ah Chung. Cheung Ping-ling helped design the poster and secured last minute sponsorship, in the hope that we would not need to utilize Department funds. Thanks also to Dean Chung for persuading the new Department Head to do the right thing by providing funds from a Department account to cover most of the shortfall that remained after the conference. The list of students who assisted during the conference is too long to reproduce here, but a special thanks is due to the four who led teams of other students to look after virtually all the details of actually running the event itself: without the unflinching assistance of Bosco (Wu Wing Keung), Candies (Lo Kwan Yuk), Coey (Hui Ka Yu), and Ringo (Cheung Siu Ko), the conference could easily have descended into chaos. I also deeply appreciated the colleagues not on the Organizing Committee, and the numerous scholars from other institutions in Hong Kong, who assisted by chairing one or more of the conference's 36 concurrent sessions. Obviously, the scholars who submitted papers, took the trouble to come all the way to Hong Kong for the conference, and especially those who re-submitted revised versions for this publication, deserve credit for actually making the conference so successful. But most of all, for the care and attention she paid to so many aspects of the planning, implementation, and aftermath of the conference (e.g., processing literally thousands of emails sent to the kantinasia gmail account), including this publication (e.g., helping to prepare the index), and for encouraging me to move forward with this long-standing dream even before Prof. Cheng did, my wife, Natalya (Lok Yuen Ching), deserves credit for the success of both the conference and its published proceedings. Together we dedicate this book to our daughter, whose energetic disposition is a source of constant good cheer, engendering ever-renewed faith in the possibility of the unity of personhood.

Stephen R. Palmquist

Hong Kong, 11 October 2010

## Note on References

References to Kant's works are normally included in the main text, specifying the volume and page numbers of the Berlin Academy Edition (or, in the case of *Critique of Pure Reason*, the standard A/B page numbering). English translations used are those of the Cambridge Edition of Kant's works, unless otherwise noted. Some authors also use the following abbreviations to refer to works named repeatedly:

<i>CPR</i> :	<i>Critique of Pure Reason</i>
<i>CPrR</i> :	<i>Critique of Practical Reason</i>
<i>CJ</i> :	<i>Critique of Judgment</i>
<i>GMM</i> or <i>Groundwork</i> :	<i>Groundwork of the Metaphysics of Morals</i>
<i>MM</i> :	<i>Metaphysics of Morals</i>
<i>RBR</i> or <i>Religion</i> :	<i>Religion within the Bounds of Bare Reason</i>

The specific volumes of *The Cambridge Edition of the Works of Immanuel Kant*, general editors Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press), listed alphabetically by title, are as follows:

*Anthropology, History, and Education*, ed. G. Zöller and R.B. Louden, tr. M. Gregor et al. (2007).

*Correspondence*, ed. and tr. A. Zweig (1999).

*Critique of the Power of Judgment*, ed. P. Guyer, tr. P. Guyer and E. Matthews (2000).

*Critique of Pure Reason*, ed and tr. P. Guyer and A.W. Wood (1998).

*Lectures on Ethics*, ed. P. Heath and J.B. Schneewind, tr. P. Heath (1997).

*Lectures on Logic*, ed. and tr. J.M. Young (1992).

*Lectures on Metaphysics*, ed. and tr. K. Ameriks and S. Naragon (1997).

*Notes and Fragments*, ed. P. Guyer, tr. C. Bowman et al. (2005).

*Opus Postumum*, ed. E. Förster, tr. E. Förster and M. Rosen (1993).

*Practical Philosophy*, ed. and tr. M.J. Gregor (1996).

*Religion and Rational Theology*, ed. and tr. A.W. Wood and G. di Giovanni (1996).

*Theoretical Philosophy, 1755-1770*, ed. and tr. D. Walford and R. Meerbote (1992).

*Theoretical Philosophy after 1781*, ed. H. Allison and P. Heath, tr. G. Hatfield, M. Friedman, H. Allison, and P. Heath (2002).

# Contents

## INTRODUCTORY ESSAYS

<i>Stephen R. Palmquist</i> Editor's Introduction .....	3
<i>Patricia Kitcher</i> <i>Keynote Essay to Book One</i> Kant's Spontaneous Thinker and (More) Spontaneous Agent .....	36
<i>Günter Wohlfart</i> <i>Keynote Essay to Book Two</i> Metacritique of Practical Reason: Back from Kant's Universalized Egocentrism via Kongzi's Moral Reciprocity and Mengzi's Compassion to Huainanzi's Reciprocal Resonance and Zhuangzi's Ethos without Ego ...	53
<i>Chung-ying Cheng</i> <i>Keynote Essay to Book Three</i> Incorporating Kantian Good Will: On Confucian <i>Ren</i> (仁) as Perfect Duty .....	74

## BOOK ONE:

### CRITICAL GROUNDWORK FOR CULTIVATING PERSONHOOD PART I Epistemological Foundations of Personhood

<i>Chong-Fuk Lau</i> 1. Self-Cognition in Transcendental Philosophy .....	99
<i>Robert Greenberg</i> 2. A Neglected Proposition of Identity .....	109
<i>Kwok-Kui Wong</i> 3. Kant and the Reality of Time .....	118
<i>Gregg Osborne</i> 4. The Active Role of the Self in Kant's First Analogy .....	129

*Andrew Brook*

5. Kant's Attack on Leibniz's and Locke's Amphibolies . . . . . 140

PART II Personhood as a Problem for Rational Metaphysics

*Julian Wuerth*

6. The First Paralogism, its Origin, and its Evolution: Kant on How the Soul Both Is and Is Not a Substance . . . . . 157

*Ulrich Fritz Wodarzik*

7. Kants Logik des Menschen – Duplizität der Subjektivität . . . 167

*Michael Thompson*

8. Antinomy of Identity . . . . . 181

*Claudia Bickmann*

9. Kant's Critical Concept of a Person: The Noumenal Sphere Grounding the Principle of Spirituality . . . . . 194

*Christine Lopes*

10. Truth, Falsehood and Dialectical Illusion: Kant's Imagination . . . . . 205

*Wolfgang Ertl*

11. Persons as Causes in Kant . . . . . 217

PART III The Role of Autonomy in Unifying Personhood

*Rainer Enskat*

12. The Cognitive Dimension of Freedom as Autonomy . . . . 233

*Makoto Suzuki*

13. Respect for Persons as the Unifying Moral Ideal . . . . . 247

*Vasil Gluchman*

14. Kant and Virtuous Action: A Case of Humanity . . . . . 256

*Adriano Naves de Brito*

15. Freedom and Value in Kant's Practical Philosophy . . . . . 265

*Courtney David Fugate*  
 16. Moral Individuality and Moral Subjectivity in Leibniz,  
 Crusius, and Kant . . . . . 273

PART IV Judgment as the Orientation of Personal Unity

*Ulrich Seeberg*  
 17. Aesthetic Judgment and the Unity of Reason . . . . . 287

*Nils Röller*  
 18. Thinking with Instruments: The Example of Kant’s  
 Compass . . . . . 300

*Bart Vandenabeele*  
 19. Common Sense and Community in Kant’s Theory of Taste 308

*Christian Helmut Wenzel*  
 20. Aesthetics and Morality in Kant and Confucius: A Second  
 Step . . . . . 321

*Eric S. Nelson*  
 21. China, Nature, and the Sublime in Kant . . . . . 333

BOOK TWO:

CULTIVATING PERSONHOOD IN POLITICS, ETHICS, AND RELIGION

PART V The Status of Persons in Applied Ethics and Law

*Ronald M. Green*  
 22. Is There a Kantian Perspective on Human Embryonic Stem  
 Cells? . . . . . 349

*Natascha Gruber*  
 23. When Is a Person a Person – When Does the “Person”  
 Begin? . . . . . 358

*Anita Ho*  
 24. Personhood and Assisted Death . . . . . 370

*Bernhard Jakl*  
 25. Human Dignity and the Innate Right to Freedom in  
 National and International Law . . . . . 382

*Peter Schröder*

26. “Irgend ein Vertrauen ... muss ... übrig bleiben”: The Idea of Trust in Kant’s Moral and Political Philosophy . . . . . 391

PART VI Persons in Politico-Cultural Community

*Günter Zöller*

27. Autocracy: Kant on the Psycho-Politics of Self-Rule . . . . 401

*Katsutoshi Kawamura*

28. Die Person als gesetzgebendes Wesen . . . . . 415

*Stijn Van Impe*

29. Kant’s Realm of Ends: A Communal Moral Practice as Locus for the Unity of Moral Personhood . . . . . 424

*Monique Castillo*

30. Kant’s Notion of Perfectibility: A Condition of World-Citizenship . . . . . 438

*Marc Rölli*

31. Person and Character in Kant’s *Anthropology from a Pragmatic Point of View* . . . . . 447

*Phil Enns*

32. Kant and the Possibility of the Religious Citizen . . . . . 455

PART VII Persons in Ethico-Religious Community

*Susan Shell*

33. Autonomy and the Unity of the Person . . . . . 465

*Martin Moors*

34. Religious Fictionalism in Kant’s Ethics of Autonomy . . . . 475

*Predrag Cicovacki*

35. Respect for Persons as Respect for the Moral Law: Nicolai Hartmann’s Reinterpretation of Kant . . . . . 485

*Aleksander Bobko*

36. The Unity of Human Personhood and the Problem of Evil 493

*Robert Gressis*

37. How To Be a Good Person Who Does Bad Things . . . . . 501

PART VIII Cultivating Personhood in Religion and Theology

*Hans Feger*

38. Kant's Idea of Autonomy as the Basis for Schelling's  
Theology of Freedom . . . . . 511

*Mohammad Raayat Jahromi*

39. Moral Theology or Theological Morality? . . . . . 523

*Chan-Goo Park*

40. Self-Knowledge and God in the Philosophy of Kant and  
Wittgenstein . . . . . 536

*Kiyoshi Himi*

41. Kant's Philosophy of Religion as the Basis for Albert  
Schweitzer's Humanitarian Awareness . . . . . 550

*Brandon Love*

42. Kant's Religious Perspective on the Human Person . . . . . 563

BOOK THREE:

EAST-WEST PERSPECTIVES ON CULTIVATING PERSONHOOD  
PART IX Mou Zongsan and Kant on Intellectual Intuition

*Wen-berng Pong*

43. Mou Zongsan's Critique of Kant's Theory of  
Self-Consciousness in the First *Critique* . . . . . 575

*Mihaela C. Fistioc*

44. Mou Zongsan and Kant on Intellectual Intuition: A  
Reconciliation . . . . . 585

*Sammy Xia-ling Xie*

45. On Kant's Duality of Human Beings . . . . . 592

*Annie Boisclair*

46. Mou Zongsan's Interpretation of the Kantian *Summum  
Bonum* in Relation to Perfect Teaching (*Yuanjiao* 圓教) . . . . . 603



*Chaehyun Chong*

47. Confucianism and Things-in-themselves (Noumena):  
Reviewing the Interpretations by Mou Zongsan and Cheng  
Chung-ying . . . . . 615

PART X Chinese Perspectives on Self-Cultivation

*A.T. Nuyen*

48. The Kantian Good Will and the Confucian Sincere Will:  
The Centrality of *Cheng* (誠, “Sincerity”) in Chinese Thought. 627

*Scott R. Stroud*

49. Desire and the Project of Moral Cultivation: Kant and  
Xunzi on the Inclinations . . . . . 639

*Mario Wenning*

50. Kant and Daoism on Nothingness . . . . . 653

*David Cummiskey*

51. Competing Conceptions of the Self in Kantian and Buddhist  
Moral Theories . . . . . 664

*Ellen Y. Zhang*

52. What Is Personhood? Kant and Huayan Buddhism . . . . . 678

PART XI Kant in Dialogue with Other Asian Traditions

*Emer O’Hagan*

53. Kant and the Buddha on Self-Knowledge . . . . . 695

*Soraj Hongladarom*

54. Kant and Vasubandhu on the “Transcendent Self” . . . . . 709

*Ruchira Majumdar*

55. Kant’s Moral Philosophy in Relation to Indian Moral  
Philosophy as Depicted in *Srimad-Bhagavad-Gita* . . . . . 715

*Takayuki Kisaka*

56. Human Personhood at the Interface between Moral Law  
and Cultural Values . . . . . 724

*Mohsen Javadi*

57. The Idea of Moral Autonomy in Kant's Ethics and its  
Rejection in Islamic Literature . . . . . 732

PART XII Kant on Asia and Asia in Kant

*Alain-Marc Rieu*

58. The Kantian Model: Confucianism and the Modern Divide 741

*Klaus-Gerd Giesen*

59. Asian Hospitality in Kant's Cosmopolitan Law . . . . . 753

*Rein Vos*

60. Doing Good or Right? Kant's Critique on Confucius . . . . 764

*Peter K. J. Park*

61. The Exclusion of Asia and Africa from the History of  
Philosophy: Is Kant Responsible? . . . . . 777

*Soo Bae Kim*

62. Menschliche Autonomie als Aufgabe – der Autonomiebegriff  
in der Geschichtsphilosophie Kants . . . . . 791

*Simon Shengjian Xie*

63. Is Kant a Western Philosopher? . . . . . 799

EPILOGUE

*Stephen R. Palmquist*

64. The Unity of Architectonic Reasoning in Kant and *I Ching* 811

- Note on Contributors . . . . . 822

- Index . . . . . 833



## INTRODUCTORY ESSAYS



## Editor's Introduction

*Stephen R. Palmquist*

The sublime, never completely attainable idea of an ethical community diminishes greatly under human hands ... But how can one expect something completely straight to be built from such crooked wood?  
(6:100, tr. Pluhar)

### 1. Kant in Asia: Keynote Essays and Epilogue

The objectives of the Kant in Asia conference were (1) to bring together in Hong Kong philosophers from around the world for a major international conference on Kant's philosophy; (2) to use Kantian philosophy as a springboard for dialogue between Asian and Western philosophers, educating all on the richness of the Chinese philosophical tradition; (3) to examine in depth the concept of personhood and its implications for education and knowledge acquisition, ethics and self-identity, religious/political community-building, and cross-cultural understanding; and (4) to contribute to Whole Person Education by exposing students as well as academics throughout Hong Kong to a wide range of top scholars from around the world. This compilation of proceedings, consisting of revised versions of two-thirds of the papers presented, demonstrates how thoroughly the objectives were realized.

This Introduction summarizes each contribution to this collection, beginning in this section with the three keynote essays that immediately follow and serve to complete the introduction both to the conference and to this book. At the conference each keynote lecture was followed by two scholarly responses; a brief summary of each response is provided here.<sup>1</sup> Sections 2–4 of this Introduction overview 63 of the 64 contrib-

---

<sup>1</sup> Due to limitations of space, the texts of these responses could not be included here. However, video recordings of each keynote session can be viewed on the

uted essays, as organized into three “Books” of 21 essays each: Book I explores themes arising primarily out of Kant’s three *Critiques*; Book II, themes relating to the politico-cultural and/or ethico-religious applications of Kant’s theories; and Book III, themes related explicitly to the interface between Kant and Asian philosophy. Each Book, in turn, is divided into four parts containing five essays each (or six, for the three parts that contain a German essay). The 64<sup>th</sup> essay, the Epilogue to this collection, is summarized at the end of this section, as it attempts to convey the idea of the whole.

The three keynote essays that follow this Introduction correspond to the collection’s three Books: Patricia Kitcher examines personhood in Kant’s Critical philosophy; Günter Wohlfart calls for a recovery of a lost Way for Western culture and society that he believes Chinese philosophy can provide better than Kant ever could; and Cheng Chung-yung offers a new model for synthesizing Kant and Chinese philosophy in an approach that attempts to preserve the strong points of each. These, like all the essays in this collection, are presented here in systematic order, rather than in the chronological order they appeared in the conference. This is partly because this book could not include all papers presented at the conference, and partly because the architectonic requirements of a published work are quite different from those of a series of oral presentations. My summaries highlight architectonic connections as well as each essay’s relation to the overall conference (and book) theme, not necessarily the main arguments of the essay itself.

Kitcher’s keynote essay corresponds to the contributed essays in Book One. She argues against Allison and Bilgrami, that the transcendental freedom introduced in *CPR* does not point necessarily to practical freedom. Kitcher claims that from the theoretical standpoint of *CPR*, the only necessity of freedom is “to characterize the strivings of reason that are the source of metaphysical error.” Allison and Bilgrami, by contrast, claim that from the *practical* standpoint, where practical reason has “primacy” over the theoretical, there is a kind of necessity that can be read back into *CPR* as unifying reason. (A resolution of this debate might be that it depends on what standpoint one adopts in using the term “necessity”.) Kitcher assumes Kant changed his mind, from seeing reason as voluntarist in *CPR* to seeing practical reason in terms of obligation and the power of choice in *CPrR*. Expressed in terms of Kant’s

theory of the *primacy* of practical reason, Kitcher's position is that practical reason has primacy and from that standpoint *is* obligatory, even though from the derivative (less spontaneous) theoretical standpoint freedom *appears* voluntarist.

Jens Timmerman and Kwan Tze-wan responded to Kitcher. After agreeing that theoretical and practical reason are very different and should not be assimilated, though they must fit together, Timmerman offered three challenges: first, Kant's appeal to the faculty of choice predates *CPrR*; second, spontaneity is not just non-rule-guidedness, but is the imposition of a special kind of rules; and third, while we are not the author of the moral law, we are the author of moral obligation, inasmuch as we freely impose the law on ourselves. Kwan raised six comments: first, free choice (or *arbitrium*) is a perennial notion, predating even *CPR*; second, Kant's table of the cognitive faculties in the Introduction to *CJ* includes neither *Wille* nor *Willkür*; third, *Wille* and *Willkür* are but different names for one and the same faculty of desire; fourth, *Willkür* implies that, as we are neither angels nor devils, we cannot renounce reason's grip; fifth, spontaneity must be contrasted with receptivity, for Kant's point is that freedom enables us to descend to sensibility (in *CPR*) or to the lower faculty of desire (in *CPrR*); sixth, morality is not guaranteed, but must be cultivated.

Wohlfart's keynote essay challenges the legitimacy of what might be called Kant's "Critical metaphysics",<sup>2</sup> in direct opposition to Book Two's essays on the ethical, political, and religious tools Kant offers for cultivating personhood. Wohlfart's "metacritique" of Kant's Critical philosophy blames the Kantian "Ego", with its self-imposed duty to universalize all maxims in order to become moral, for being the root of the corruption and evil that besets contemporary Western culture (especially its ethics, politics, and religion). In place of this unworkable and ultimately illogical ideal, Wohlfart urges us to go "back" to the East, where a refreshing "ethos without morality" can be found. While his sketch of four key ancient Chinese philosophers may be brief, it effectively challenges "Kantian believers" to rethink the basis of their Critical faith. The question this keynote essay poses for the other essays, especially those in Book Two, is: *does* Kantian philosophy have the resources for constructing a genuinely humane and *sympathetic* culture,

---

2 See my book, *Kant's System of Perspectives: An architectonic interpretation of the Critical philosophy* (Lanham: University Press of America, 1993), X.1.



one that incorporates key insights from Asian philosophy, as well as those familiar ones from the Western tradition?

Chad Hansen and Stephen Palmquist responded to Wohlfart. Hansen argued that the idea of a “metacritique” is too romantic, over-emphasizing feeling in contrast to reason. The Chinese tradition does not consistently reject reasoning in favor of feeling or intuition. Jesus’ golden rule and Confucius’ “silver” rule both employ a two-person approach to morality, both contrasting with Kant’s. Various Mohist and Daoist concepts of morality illustrate that some Chinese philosophers reject heteronomous reasoning without going to the extreme of universalism. Zhuangzi’s position is that each person decides which *dao* will be the grounding of one’s morality, with the heart-mind enabling “*shih-fei*” (indexical, “wrong-right”) judgments as one follows one’s self-chosen *dao*. Palmquist gave eight reasons for claiming Wohlfart agrees with Kant: Kant’s is a one-off *self*-critique, whereas “metacritique” is *other*-critique; the Kantian ego is a *necessary limit* that reveals Descartes’ “Egod” to be illusory; humans are “crooked wood” that cannot be saved by a universalized will; universalization is a search for practical contradiction, not logical contradiction, as we seek to realize the holiness of genuine personhood; acting *as if* God exists is not hypocrisy, but a transcendental call to humble action in the face of empirical ignorance; “moral purism” leads to “moral terrorism” only when we portray *impure* cultural principles as universal truths; Kantian moral religion is “bare”, not “pure”; bare reason needs to be “dressed” with cultural norms, yet these must not be universalized.

Cheng’s keynote essay to Book Three provides the groundwork for a thoroughgoing synthesis between Kantian and Confucian philosophies. In a portion of the essay not presented here,<sup>3</sup> Cheng conducts an in-depth analysis of the moral philosophy in *GMM*, arguing that Kant’s distinction between perfect and imperfect duties ought not to be applied as rigidly as Kant did. Employing Confucian methods of analysis, grounded on the principle of *ren* (benevolence), Cheng argues that the “life principle” is the ultimate root of all four types of duty; when Kant’s resistance to the influence of inclinations is moderated by a Con-

---

3 Professor Cheng’s original essay dealt with many of the intricacies of Kantian philosophy and how they can be transformed by a Confucian re-interpretation. Due to length limitations, the first and longer part has been omitted here and is due to be published in an upcoming special issue of the *Journal of Chinese Philosophy* (2011).

fucian awareness that human needs can be good as well as bad, Kant's perfect duties turn out to appear "imperfect" in some life situations, while imperfect duties can become perfect duties. The portion of Cheng's essay presented here focuses on how the Confucian doctrine of *ren*, when interpreted in a Kantian framework, serves as the necessary groundwork for a Sino-Kantian understanding of human personhood that puts a human face on Kantian formalism in ethics, while adding rigor to Confucian moral philosophy.

Hans Feger and Philip J. Ivanhoe responded to Cheng. Feger noted that Kant's moral theory is intentionally ideal, whereas human beings are fallible. *CPrR* is not about moral behavior but about the conditions necessary for *any* moral behavior to be possible. Unlike the golden rule, the categorical imperative cannot be applied directly to specific cases. We learn about freedom from our experience of the moral law, yet freedom is the transcendental essence of morality and is therefore itself unknowable. While the Confucian tradition focuses on empirical morality, Kant presents the limits of all genuine morality, thereby undermining yet potentially grounding *any* moral philosophy (such as the Confucian) that prescribes specific moral goals. Ivanhoe compared Cheng's critique of Kant to Nancy Sherman's attempt to naturalize Kantian ethics from an Aristotelian perspective, though Cheng's attempt to "rescue" Kant leans more toward the "onto-cosmological" tradition reminiscent of Mou Zhongsan. Cheng's use of *ren* is problematic: assuming *ren* as a mandate of Heaven seems more like sanctification than naturalization. Rather than rescuing Kant's transcendental project, it replaces the search for a transcendental foundation with a Heavenly ordained natural order. This is one plausible way to read the Confucian texts, but it may not be the most promising way to construct contemporary ethical theory or to build upon Kant's remarkable insights.

Following the three Books, consisting of 63 essays (60 in English and three in German), this collection concludes with a revised version of my conference paper, suggesting the conference theme can best be appreciated by relating Kant's "architectonic" to the oldest book of Chinese philosophy, the *I Ching* or *Book of Changes*. When Kant emphasized the need to grasp the "idea of the whole" (Bxxx) to understand his philosophy, he was alluding to the categories as formally structuring his architectonic plan. Understanding this crucial connection, as mediated by the "I think", enables us to appreciate why he thought architectonic reasoning is required not only for all genuinely philosophical thought, but also for the unity of human personhood (the conference's

sub-theme). Intriguing formal similarities exist between Kant's application of categories and the *I Ching's* application of 64 hexagrams. Adopting the standpoint of the *I Ching*—something Kant himself was probably too “Western” ever to have done!—I seek guidance on how this very connection (of Kant with Asian philosophy and Asian philosophy with Kant) is to be understood. As if to confirm Xie's claim in Chapter 63 (see §4, below), the *I Ching* conveys a surprisingly relevant insight.

## 2. Critical Groundwork for Cultivating Personhood

The idea uniting the twenty-one diverse essays in Book One is that Kant's three *Critiques* offer a systematic account of what it means to be a human person, *in abstracto*. The four parts follow the order of the three *Critiques*, with the first two parts devoted to *CPR's* first half (through the end of the *Analytic*) and second half (from the *Dialectic* onwards), respectively. The Kantian picture of the human person emerges from Book One as paradoxically both divided and united when considered in abstraction, depending on which Critical standpoints one adopts. This suggests we can and must realize our personhood *only* through the social relations that constitute human communities. What is often neglected, especially by superficial overviews of Kant's philosophy, is that this *is* the new “moral metaphysics” that the three *Critiques* served to ground. This is why Book Two covers themes relating to the cultivation of personhood through human communities. Book Three's emphasis on Asian philosophy demonstrates that this new emphasis Kant brought to Western metaphysics is standard procedure in the East.

Part I focuses on topics from the first half of *CPR*, beginning with Lau Chong-fuk's essay examining what Kant means by “transcendental”, especially in *CPR's* Prefaces. Lau distinguishes between empirical and transcendental self-cognition, arguing that Kant does not intend the distinction between the faculties of sensibility and understanding (together with their respective formal conditions of spatio-temporality and categoriality) to be psychological. Transcendental cognition is not about our cognitive mechanisms, but refers to a new type of conceptual analysis of what it means for finite rational beings to have knowledge. Even the “I” of pure apperception is not something we ever experience; transcendental cognition is therefore neither phenomenal nor noume-

nal, but concerns only the necessary preconditions for phenomenal knowledge and noumenal action. *CPR's* theory of human nature is bound to be abstract, since it concerns only what we must assume in order for self-cognition to be possible. However, as we shall see, this does not prevent it from having specific psychological implications, once a metaphysical structure is built on this transcendental groundwork.

A crucial problem faced by anyone hoping to locate a Kantian theory of the unity of human personhood is that Kant appears to adopt two very different approaches to explaining what it means to be an object of knowledge—a challenge effectively met by Robert Greenberg's essay on a neglected "identity proposition" Kant advances immediately following *CPR's* Introduction. Kant's proposition unites his apparently conflicting foundationalist and causal theories of object causation, suggesting that these two theories of knowledge are somehow identical. Greenberg's intriguing claim, that "the double-object interpretation can be derived from the dual-aspect interpretation", is sure to spark new debate on the proper interpretation of transcendental idealism. This would explain why *both* theories of knowledge appear intermingled, throughout *CPR*.<sup>4</sup>

That *time* poses special problems for Kant's understanding of the unity of human personhood is highlighted in Wong Kwok-kui's essay on the various roles time plays throughout *CPR's* Aesthetic and Analytic. Wong investigates whether any aspect of Kant's theory can provide an effective response to St. Augustine's pondering on the reality of time as an extensive magnitude. The transcendental ideality of time in the Aesthetic is not of much help here, since time there functions merely as a form of intuition. More help is gained in the Analogies, where time is a necessary presupposition of our experiences of permanence, succession, and co-existence. The crucial passage, Wong argues, is the A-Deduction's account of the threefold synthesis of imagination, for here Kant "confront[s] the problem that 'nothing abides' in the flux of time", arguing that time must be a component in any understanding of what constitutes a "unit of knowledge".

---

4 For a similar explanation of the limitations of Allison's interpretation, insofar as it neglects texts that clearly support what Greenberg calls Kant's causal theory, see *Kant's System of Perspectives*, apx.VI (esp. 393–4). However, I did not there call attention to the proposition Greenberg highlights.

Gregg Osborne further examines Kant's explanation in the First Analogy of how we can so much as believe that objects come into being and cease to exist. While the abstract arguments considered here might seem far removed from anything constitutive of human personhood, and Osborne himself remains unsure that his reconstruction of Kant's arguments succeeds, what clearly emerges from the analysis is that such beliefs are grounded in functions of judgment, and so also, of the self. In the arguments under consideration Kant distinguishes between what is objective and what is merely subjective in our experiences of temporal relations. Understanding the grounds for such a crucial distinction as that between objective and subjective coming into (or passing out of) existence clearly lies at the heart of any theoretical account of human subjectivity.

Andrew Brook concludes Part I by examining the similarly rarefied arguments Kant develops in the Amphiboly, especially his emphasis on "transcendental reflection" and "numerical identity". Brook ponders why Kant waits until an appendix to the *Analytic* to introduce the former and discuss the latter, even excluding identity from the table of categories. Not satisfied with the explanation that identity is a principle of General Logic, defining logical consistency, not a transcendental principle that helps produce objectivity (as, for example, causality is), Brook argues that the omission of identity renders suspect any claim to the alleged completeness of Kant's table of categories. Less controversially, he observes that Kant's best (and almost his only) arguments that all our experience must be grounded in sensible intuition also appear in this appendix. Introducing transcendental reflection at this turning point, it seems, serves to focus the reader's attention on the importance of identifying the source *for the human person* of anything we believe constitutes objective knowledge.

Part II (topics covering *CPR*'s second half) begins with Julian Wuerth's essay on the First Paralogism, arguing that Kant's criticism of rational psychology here applies only to the attribution of permanence (and hence, immortality) to the soul, not to the soul's nature as a substance or thing in itself *per se*. Quoting a wide range of texts from throughout Kant's *corpus*, Wuerth demonstrates how the criticism of traditional metaphysics in the First Paralogism mirrors and applies the insights introduced in the Analogies and the Amphiboly (the main section of *CPR* that precedes the Paralogisms). Whereas phenomenal substance (defended in the First Analogy) implies permanence, noumenal substance does not; yet this does not make the latter ontologically vac-

uous. Failing to engage in transcendental reflection, the rational psychologists conflated phenomenal and noumenal substance and illegitimately inferred the soul's permanence. Kant's conception of human personhood comes out looking more robustly metaphysical on Wuerth's reading than many accounts of the Dialectic have assumed.

Ulrich Wodarzik's essay (in German) follows with an account of how Kant's transcendental logic constitutes an advance on Aristotle's bivalent (I-it) logic. The central feature of Kant's trivalent logic is the "double I" arising from the phenomenal-noumenal distinction: the logical or transcendental "I-1" must (as shown in the Deduction) be assumed as spectator, while the empirical "I-2" can (as argued in the Paralogisms) be perceived as an object (an "it") like other phenomena. A third "I" arises out of the distinction between "I-2" and the phenomenal "it": this "I" recognizes the presence of a "you" in the world. Wodarzik argues that freedom and morality, and so also the unity of human personhood, arise out of this "I-You" distinction, for it sets up a transcendental principle of intersubjectivity that requires us to respect other persons by acknowledging them as free, noumenal beings. Whereas Aristotle's theory is unable to distinguish between the not-I as a thing and the not-I as a person, Kant unifies these in a threefold theory of personhood.

Michael Thompson calls attention to an "Antinomy of Identity" hidden in the conflicting discussions of personal identity advanced by empiricists and rationalists prior to Kant, and mirrored in the Confucian and Daoist traditions. Displaying a typical pair of antinomical arguments, Thompson "synthesize[s] the two positions into a unity that affirms human essence, embraces personal identity, and celebrates contingency." Whereas the Locke-Confucius defense of the antithesis requires the problematic assumption that a person who experiences a sudden change in life circumstances actually has a change in personal identity, the Kant-LaoTzu defense of the thesis (in terms of a transcendental "I" or a "greater identity, the Tao") problematically assumes "we are never conscious of this so-called unity of consciousness." The former overlooks that we experience ourselves as phenomenally the same; the latter "smacks of empty formalism". Thompson defends "the fact of identity", by examining the phenomenological life "context" of human beings, appealing to our temporal, bodily nature and to the orientation to the past and future that consciousness gives us.

The foregoing essays on the Paralogisms and a new Antinomy are followed by Claudia Bickmann's essay, viewing all three ideas of reason

from the theological standpoint of the Ideal. She argues that, interpreted in light of the unity of apperception, Kant's three ideas of reason constitute the *spiritual* orientation of Kant's theory of human personhood. It is spiritual because Kant, like Hume and Locke, seeks to give due place to the "ever-floating, never-resting" character of our phenomenal experiences, yet insists on grounding this in an underlying transcendental reality. Kant's distinctiveness is to insist that this reality remains unknowable. The transcendental ideas of our soul, the world, and God, when interpreted heuristically as reason's tools for obtaining systematic unity *within* our otherwise disparate experience, serve as a "quasi-schema" to give us confidence even though we cannot actually appeal to knowledge as such. This approach, Bickmann observes in opening and closing, is akin to Hindu Advaita-Vedanta philosophy, especially as revised by the Buddhist and Taoist critiques of an abiding self.

Christine Lopes sheds light on the general status of the ideas of reason by examining the impact Kant's theory of imagination has on both his epistemology and his criticism of metaphysics. In contrast to interpreters who see Kantian intuition as essentially conceptual, Lopes argues intuition must contain a "non-fully-fledged conceptual" element. This element gives rise to the possibility of *error* in our empirical judgments, as imagination links intuitions with concepts. It is also the key factor giving rise to the transcendental illusion Kant analyzes in the Dialectic, for such illusion is also a species of error, whereby we mistakenly identify "merely formal acts of conjecturing and reasoning about the possibility of objects in general" as "substantive mental acts (i. e., acts of cognition of real objects)". Such illusion arises and is inevitable as a direct result of the non-fully-fledged conceptual element in intuition; on this basis human imagination is able to fool itself into mistaking transcendent for empirical objects.

Wolfgang Ertl links Part II's focus on the Dialectic and Part III's focus on the second *Critique* by arguing that Kant has a solution to the freedom-determinism problem that goes further than Strawsonian compatibilism, avoiding the extremes of reason's causal inertness and causal overdetermination. Kant regards persons (given their nature as substances, called souls) as having a real, causal influence on the world through free choices. But this, Ertl argues, is not a separate kind of causality, entirely unrelated to the natural causality that Kant defends in the Analogies; rather, they are one and the same causal relation, viewed from either the practical or theoretical standpoint, respectively. While he admits to leaving a number of key questions unanswered, Ertl pro-

vides a fruitful basis for a deeper understanding of the theoretical aspects of Kantian personhood.

With Ertl's preparation complete, Part III examines autonomy, the core of Kant's theory of human personhood. Rainer Enskat sets the tone by arguing that the doctrine of autonomy carries with it a specific (and often neglected) "cognitive dimension". Kantian autonomy is a twofold faculty, consisting "*primarily* in the *cognitive faculty* to judge and to recognize and, secondarily, in the *practical faculty* of so-called causality of the free will, to exercise ... morally-consistent ways of acting." Reconstructing the key formal steps in Kant's ethical theory, Enskat focuses on the specific case of lying. Kant's anthropological writings portray human beings' tendency to lie in a less formalistic, highly pragmatic manner. These two sides of Kant's views on lying and truth-telling do not contradict, but correlate nicely together. The practical unity of human personhood, Enskat maintains, is grounded in our cognition of this (reconstructed) formal judging procedure.

Whereas Enskat focuses on an abstract formalization of Kant's ethics based on the first formulation of the categorical imperative (the universality of moral maxims), the next two essays focus on the second formulation, respect for persons. Makoto Suzuki argues, along the lines of David Ross, that common sense moral judgments show a partiality that seems difficult to reconcile with Kantian respect for persons. After examining the attempts to rescue Kant made by several prominent Kantian ethicists, Suzuki concludes: "The idea of respect for persons is *not by itself* a unifying moral ideal; some additional factor, or some distinct or more fundamental ideal is required for making sense of the thought that commonsensical requirements are moral and true duties."

As if responding to Suzuki's challenge, Vasil Gluchman claims Kant's appeal to "humanity", as a principle for guiding moral decisions, is antithetical to common sense morality. Kantian morality attempts to overcome the "crudity" of the human nature we hold in common (e. g., our tendency to lie), as exemplified by Kant's call to treat strangers as equals to one's friends; yet Gluchman claims that "humanity", like "animality" for our mammalian cousins, can ground us in our specific (common sense) preferences. He analyses "humanity" as having a twofold meaning and application: as a pure moral concept, and as an "added value". He concludes with an urgent call to both individuals and societies to take on board this principle in shaping our future life together. This provides a foretaste of the essays in Book Two, where the role of the community in cultivating personhood is examined more fully. Both



Suzuki and Gluchman hint that the Kantian notion of *dignity* might serve as a much-needed supplement to the basic principle of Kantian respect for persons.

This hint becomes the focus of Adriano de Brito's essay, objecting to the use compatibilists often make of Kant: our intrinsic value, not our freedom, is what makes us human beings responsible for our moral actions. Personhood therefore depends more on dignity (i.e., on an agent's value) than on whether one's choices are technically free or determined. Kant's theoretical concept of freedom is, at best, negative; it is a quasi-Humean expression of "ignorance and uncertainty about the causes of a practical phenomenon". Citing (like Enskat) the example of lying, de Brito demonstrates how the faulty assumption, that human freedom is the justification for assigning responsibility to human actions, is the basis for the "obscene" claim that Kant's moral philosophy makes evil actions heteronomous, thereby removing responsibility from evildoers. The only way to avoid this undesirable conclusion is to identify the "fact" of freedom with our actual moral judgments that persons *are* responsible (i.e., in our determination to value them as such), not in any factual knowledge that we are empirically free.

Courtney Fugate concludes Part III, exploring how the concepts of individuality and subjectivity interrelate in the moral theories of Kant and his predecessors, Leibniz and Crusius. While Kant's critical method renders individuality and subjectivity compatible, it provides no justification for their *unity*. Crusius had argued against Leibniz's pure moral individualism, claiming that "an absolutely self-determining or causally self-transparent subjectivity" must lie at its basis. By synthesizing these positions, "Kant arrives at the equivalence of a free will and a will governed by the categorical imperative". While retaining Crusius' absolute freedom, Kant shows that "this same absoluteness precludes its taking on an externally-given form" even though "its nature as will requires it nevertheless to have a form intrinsic to it." Although Fugate is doubtful about Kant's success, he portrays Kant as aiming to avoid the difficulties of these previous metaphysicians by casting "the seemingly incompatible positions of moral subjectivism and moral individualism as dynamic moments in a self-developing and internally complex teleology of moral self-consciousness."

Part IV completes Book One with essays covering themes corresponding to what I have elsewhere called the "judicial standpoint" of

Kant's third *Critique* (*CJ*).<sup>5</sup> Ulrich Seeberg's comprehensive overview of the whole Critical philosophy claims the idea of the intelligible world, including *both* "the sensible world (nature) and the subject as being aware of the sensible world", is the key to Kant's understanding of the unity of human personhood. In *CPR* the unity of apperception and the unity of nature are directly correlated: both are unknowable in themselves, yet presupposed by any objective knowledge. The fact of moral obligation arises directly out of this unified picture of subject and object jointly belonging to an intelligible world, for our ignorance of the true causes of our actions makes us subject to *duties*, virtuous rather than holy. Aesthetic judgment thus makes constitutive (for the feeling of the beautiful in nature) what for practical reason was a mere "factum": an intelligible reality *must* exist, otherwise our experiences of purposiveness would be meaningless. Despite Kant's affirmation of human selfhood and a divine being, Seeberg finds these suggestions tantalizingly similar to certain Eastern philosophies.

Nils Röller's essay on Kant's metaphorical use of the compass aptly illustrates how judging reason orients every aspect of the Critical philosophy, not only his theories of beauty, sublimity, and purposive organisms. Kant's interest in science, for example, was focused on imaginative speculation rather than experimental rigor. Kant's fondness for the compass as a metaphor for the orientation provided by Critical philosophy relates to its function as an instrument (a human invention) that detects real forces, external to ourselves, yet nevertheless guides and orients our exploration of the world. For us humans, reason must also serve as a tool that puts us in touch, through rational faith, with a reality that is for us unknown, yet can be detected by those who possess the instrument.<sup>6</sup>

Bart Vandenabeele examines the significance of *CJ*'s theory of taste as emphasizing the need for "community with others who share our sensibilities and capacities to judge the beauty of nature and art". Interacting with a range of contemporary interpreters (especially Allison), Vandenabeele argues that Kant's concept of "universal communicability" makes sense only if we recognize its reciprocal relationship with the disinterestedness of the pleasure we feel in making aesthetic judgments: disinterestedness "is the essential, *a priori* condition for the universal

---

5 See *Kant's System of Perspectives*, II.4 and IX.1.

6 I argue in *Kant's System of Perspectives* that Kant's compass is his "architectonic" art; I explore its connection with Chinese ("orient-al") ways of thinking in essay 64.

communicability of aesthetic judgments”, while one’s certainty of actually being disinterested is measured by the judgment’s universal communicability. After examining Kant’s reasons for thinking aesthetic judgments must appeal to a “common sense”, Vandenaabee clarifies that such judgments are important not so much because they “make a transition from the sensible to the supersensible,” but because they enable us to cultivate personhood by “rethinking subjectivity and intersubjectivity as manifest within felt, particularized pleasures that are universally shareable—not despite but *due to* their affective nature.”

Christian Wenzel’s essay is the second in a series comparing and contrasting Kant and Confucius on beauty.<sup>7</sup> Because Kant strictly separates morality from aesthetics, the pleasure in the good and the pleasure in the beautiful must be of two different kinds, and moral acts themselves cannot be beautiful. Nevertheless, Kant indicates possible connections between morality and beauty in his *CJ* comments on aesthetic ideas, symbolism, the *sensus communis*, and education. Confucius, by contrast, postulates no such radical separation between beauty and morality, but sees the *dao* as interpenetrating both and expressed in ritual: moral acts can be beautiful. One might wonder whether Confucius simply missed the crucial difference here, or whether Kant overemphasized the separation. Wenzel argues that, in spite of the fundamental differences between their theories, traces of transcendental idealism can be seen in the Confucian appeal to *dao* and *tian*.

Eric Nelson’s essay appropriately transitions to both Book Two’s community focus and Book Three’s Asia focus. Placing Kant’s “anthropological speculations concerning the Chinese” in their “dubious” context “of Enlightenment discourses about race”, he notes the distinctive “depictions of the grotesque and the sublime and of absorption into the inhuman” in Kant’s pre-critical depictions of Chinese culture. Focusing on *CJ*’s references to “the feeling of life”, Nelson rejects the allegation in Wohlfart’s keynote essay, that Kant endorses mankind’s domination of nature. Rather, Nelson’s reading of *CJ* suggests “a middle ground between impersonal nature and moral personality” for both Kant and Daoism. Contrary to the Confucian and Kantian critiques, that Daoism “dissolves the human into mystical nature”, the classical Daoist texts focus on nature’s ethical function in cultivating the sage. The playfulness entailed by such cultivation is not unlike the “free play” promoted in

---

7 For the first essay in the series, see Christian Helmut Wenzel, “Beauty in Kant and Confucius: A First Step”, *Journal of Chinese Philosophy* 33 (2006), 95–108.

CJ. Nelson offers a refreshing alternative to the all-too-frequent claims that “Kant privileged the human and neglected the natural”, while “early Daoism neglected the human in prioritizing nature.” Kant and Daoism both promote “individuation through cultivating balance in relation to nature within and outside oneself.”

### 3. Cultivating Personhood in Ethical, Political, and Religious Community

Book Two's coverage of the manifold ways Kant believed personhood can be cultivated in community begins in Part V with three Kantian responses to contemporary issues in applied ethics. Ronald Green examines two conflicting ways Kantians might understand our obligations to prenatal forms of human life. The first appeals to the second formulation of the categorical imperative, noting that Kant attributes “moral personhood only to beings capable of actively exercising reason.” The second highlights Kant's “extended sense” of this rationality criterion, applying it to those who are only potentially rational. This second approach claims we have a duty to respect very young children, and arguably also preborn children, because not to do so would develop *in us* unhealthy habits that could lead us to abuse older, rationally developing, children: “Any cultivation of behaviors and attitudes of violence toward children threatens the safety and well being of all children.” Green suggests shifting the focus “away from uncertain questions about which properties must be possessed by incipient human beings to qualify them as persons” and to “the implications of mistreating a class of human beings”. Destroying embryos “does not tug at our heartstrings”, Green argues, “nor will it produce psychologically or physically damaged adults.”

After providing a detailed technical account of human embryonic stem cell research, Natascha Gruber argues that Kant's ethics cannot possibly be used to solve this debate, one way or the other. Gruber develops her argument as a response to Manninen: “*if* empirical (‘physical’) operations cannot causally render the creation of a free being, then *nowhere* along the line of the development of a human being can the coming into ‘existence’ of reason and freedom be pointed out, because these developmental stages from zygote, blastocyst, embryo, infant, and so on, are all empirical or biological, phenomena.” Gruber's

claim is that, if we understand Kant's phenomenal-noumenal distinction as proposing a two-perspective view of these two realms, rather than as involving causal relations between them, there can never be *any* link between the physical processes of our biological development and the practical rationality that gives human beings dignity. Kant's use of "*herübergezogen*" indicates how he believed personhood begins: "*via* the act of procreation a being is drawn *from* the 'intelligible' *into* the phenomenal world."

Anita Ho continues this applied ethics theme by arguing that, contrary to the way he is often portrayed, Kant's writings on personhood and suicide do not rule out all forms of euthanasia and altruistic suicide. While Kant does not allow rational agents who are experiencing irreversible *physical* decline to will the destruction of their personhood by requesting assisted death, he ironically allows euthanasia in situations where patients are losing their *rational* capacity, regardless of their physical condition. Ho concludes by suggesting that Confucian values, with their family-oriented emphasis, provide a stronger basis for rejecting attempts to justify such forms of killing as ethical: if our personhood lies not in our *individual* rational capacity but in our *family relationships*, then mental or other serious forms of illness do not make us non-persons.

Focusing on the political side of Kant's ethics, Part V continues with Bernhard Jakl's essay examining the German Constitution's appeal to human dignity as the fundamental principle of the entire legal system. Recent debates in Germany, especially regarding a proposed law that would permit the shooting down of an airplane that had been hijacked for terrorist purposes, have focused on the role of human dignity in law. Jakl draws attention to the key distinction between the foundational application of the categorical imperative in Kant's *Groundwork* and its legal application in the *Metaphysics of Morals*, Doctrine of Right. The issue of human dignity arises primarily in the former, ethical and foundational context. In the latter, properly legal context Kant appeals instead to the one and only innate right each human person has: the right to be a free agent. This, Jakl argues, ought to be the focus of attention in legal debates, because this is the "specifically legal" meaning of dignity for Kant.

Peter Schröder concludes Part V by discussing the role of trust in counter-balancing humanity's conflictual nature. Kant's denial of any "*right to lie*" relates to *legality*, not interpersonal ethics. The *duty* to be truthful is imperfect, so we must trust people in private. But truth-telling in public contexts is a *perfect* duty: one injured by a public lie has a

right to take legal action. Citizens do not have a *right* to revolt; political community is possible only if people trust the sovereign, who may coerce them to obey national laws. In international relations, trust becomes even more significant because no sovereign exists to limit the freedom of individual states. Kant thus compares relations between states to interpersonal relationships: trust is more important in international relations than in ordinary political relations (where “right” makes trust secondary), because states, like individuals, cannot coerce another state/person who lies. *Perpetual Peace* takes a “transcendental” turn by requiring even warring states not to engage in activities that would make subsequent *trust* impossible.

Developing cultural applications of this political theme, Part VI explores Kant's account of how personhood is cultivated in legislative *communities*. Günter Zöller explores Plato's and Aristotle's influences on Kant's ethical and political philosophy: while Kant follows their division of practical philosophy into ethics and politics, systematically using political concepts to structure ethical discourse, he departs from them in assigning *a priori* knowledge to the practical realm, and unifies them through a shared dependence on the categorical imperative. Kant “cast the origin and status of the moral law in terms of the political institution of state legislation”. Citing Kant's introductions to the *Metaphysics of Morals*, Zöller shows how key legal/political concepts, such as “legislation” and “constraint”, inform Kant's presentation of the relational structure of the practical self. Kant chooses a political term, “autocracy”, to distinguish specifically *ethical* autonomy from “autonomy” in general: we *govern* feelings and desires, just as political systems govern societies. Kant's practical philosophy emerges as a comprehensive synthesis of ethics and politics. Two types of legislation apply law either internally, by referring to ethical motivations, or externally, by refraining from doing so (as in politics).

Katsutoshi Kawamura's essay (in German) examines the role of self-legislation in Kant's *Groundwork*, asking whether it pre-exists as a condition of the human will or must be created by moral agents. Kant definitely affirms that the moral law's legislation is *a priori*: it takes place at a timeless, noumenal level, before actual moral maxims are formed (cf. Part III). However, this does not mean autonomous self-legislation is the “default” situation for us, otherwise why would we so often act contrary to the moral law? Kawamura suggests instead that at the phenomenal level autonomy exists as a *potential*: each person must activate it by learning to make self-legislated choices. This is where “respect”

comes into play in Kant's moral theory: we respect the moral law each time we strive to conform our (phenomenal) moral maxims to the (noumenal) standards of morality. In this sense Kantian morality presents itself as "a task".

Viewing Kantian morality as a task requires us to recognize ourselves as part of an ethically-grounded community, as Stijn Van Impe shows in his thoroughgoing analysis of Kant's theory of the "realm/kingdom of ends". After examining Kant's various definitions of this key term, van Impe unveils its multi-layered meaning, referring "to the union of rational lawgiving beings as ends in themselves and the unification of their morally permissible subjective ends as well as of their guiding maxims that moral actions ought to be based upon." He further analyzes how the realm of ends functions as a practical "ideal" or normative "archetype", grounding communal moral practice in a way that brings about the "complete determination" of the categorical imperative. Emphasizing the social and communal dimensions of Kant's realm of ends, Van Impe views this multi-layered ideal "as the locus for a communal moral practice based on the unity of moral personhood".

In Kant's philosophy of culture, Monique Castillo argues, cosmopolitanism and perfectibility link political philosophy to anthropology. Globalization typically focuses on *technical* advances, but Kant's concept of cultural advancement emphasizes *moral* evolution, requiring human solidarity in the mutual search for perfectibility. As *nobody* "owns" one's personhood, this potential fulfillment of human destiny must be cultivated in cooperation with others. Castillo highlights two related theses Kant defends in *CJ*: "nature acts against my goals"; and "Culture claims an ethical and non-technical answer to the relationship between freedom and nature". The first provides anthropological evidence that life is meaningful, rooting the fulfillment of our self-set goals (i. e., happiness) in something we cannot control (nature), but tempts us to use technical means to combat nature's opposition. The second reminds us that cultural creations (i. e., rational symbols) provide the only meaningful protection against this temptation, for they enable us to respond ethically (*from within*) to the "denaturation" we face in relating to nature. We therefore realize perfectibility only in relationships with *other* cultures and generations.

Marc Rölli assesses whether the theory of culture in Kant's *Anthropology* is consistent with his universal ethics. For Kant "character" refers both to a person's moral "essence" and to the empirical details that manifest it. Unlike our natural character (i. e., aptitudes and temperament/

disposition), our true character relates to our moral nature and therefore “has to be acquired and cannot be given by nature”. Being empirically conditioned, dependent on concrete observations of real tendencies among human beings, yet appealing to something quasi-universal (i. e., ingrained in a person through an attunement of nature and culture), the concept of character is a “deeply ambivalent” means of connecting the “psychological, medical, racial, and ethnic aspects [of human beings] with each other.” The changes in Kant’s theory during his lifetime, Röllli concludes, were not radical shifts toward or away from a universal conception of humanity. Rather, Kant’s anthropology exhibits a hierarchical structure whereby his universal moral theory serves as the “transcendental benchmark” for empirical distinctions relating to human character.

Part VI concludes with Phil Enns’ essay, distinguishing Kant’s view of the way laws achieve their lawfulness from that of Habermas and Rawls. For the latter, laws must not only pass Kant’s rational test (i. e., universal applicability under the principle of right), but their matter must also be justified independently. Kant’s approach allows “religious reasons [to] play a significant role in political deliberation and justifying political decisions.” In private reasoning, officers or employees of institutions (e. g., government or religious officials) must *obey the rules*. But in public, when not bound by commitments to loyalty, people are free to criticize the rules. Societies are free insofar as they welcome criticism by those not charged with the duty to obey. Public uses of reason shape the rules as societies develop and mature. Kant, unlike Habermas and Rawls, thus leaves religious people free to participate in public debate on the relevance of religious ideas to the political sphere *and* about how religious rules may need to change.

Following Enns’ lead, Part VII focuses more explicitly on the interface between ethics and religion, beginning with Susan Shell’s essay on a paradox Kant thinks every moral being faces: the moral law imposes a constraint that we both want (as moral beings) and do not want (as animal beings) to obey. This paradox is rooted in and reveals our dual nature, as both noumenal and phenomenal beings, and poses the problem of the unity of human personhood from the moral (or practical) standpoint. After examining how Kant modifies the way he dealt with this problem in *CPrR* (as compared to the earlier *GMM*), Shell explains why Kant’s new stance on the nature of human autonomy requires an appeal to religion—a peculiar religion that focuses on an internal



lawgiver who is “beyond human comprehension without requiring our submission to anything external to ourselves.”

Martin Moors examines the role of the “as if” in Kant’s interpretation of religion, arguing it is as crucial in *Religion* as in the Dialectic of *CPR*. The philosopher’s use of “as if” reflects a “suspicion of irreality” in religious utterances; yet it also requires reflecting in a manner that commits oneself (subjectively) to the truth of various religious claims, for the purpose of enhancing one’s moral empowerment. Moors examines three examples to demonstrate how Kant’s reasoning works: the postulate of a “divine Obligator” in the Dialectic of *CPrR*; the appeal to the “Son of God” in the Second Piece of *Religion*; and the definition of religion as “the recognition of all our duties as divine commands” to legitimize belief in God as a beneficent world ruler. This “fictionalism” in Kant’s understanding of religion does not deny the truth of such claims; rather, it views their truth from a new standpoint, focused not on objective verifiability, but on the importance of such beliefs for “the entire orientation of my life”.

Pedrag Cicovacki shows how Kant’s dualistic/Christian conception of human personhood, with its three levels of “predisposition”, relates to the moral qualities that make persons worthy of respect. While our animality does not directly require respect for persons, and personality clearly does, the middle predisposition (“humanity”) is more controversial. Cicovacki’s central question is: must we respect *all* persons equally? He distinguishes negative respect (equally deserved by all, even immoral, persons) and positive respect (deserved only by the virtuous), urging us to broaden Kantian respect by including *all* aspects of human nature, even those relating to animality. Cicovacki employs Nicolai Hartmann’s reformulation of the categorical imperative to effect such a broadening. Hartmann’s principle, implied by Kant’s formulations, is that moral choices must never neglect another person’s individuality while imposing the universal “ought”. Hartmann thus challenges Kant’s exclusion of feelings such as love from the realm of moral value. Love can unite persons in community more effectively than purely rational values such as justice.

Despite the emphasis of Kantian ethics on universal law as necessary to transform reality “into a rationally ordered unity”, Aleksander Bobko argues that Kant’s concern for “personal development” eventually leads him to acknowledge the “helplessness of virtue”. Reason thus becomes as dialectical in its practical manifestation as in its theoretical form. Whereas reason’s theoretical antinomies “pose only intellectual prob-

lems”, its practical antinomy threatens to “introduce a dangerous kind of chaos into the life of the individual and of the entire human community.” Because of the radical evil in human nature, “persons of good will cannot be sure their rational actions will not yield adverse effects that render a service to evil; evil in an incomprehensible way destroys the order established by the moral law.” As such, “evil disturbs the unity of human personhood.” The only solution is for people of good will to adopt a mutually agreed “unifying principle”, uniting in an *ethical* community through belief in God.

With Bobko having drawn attention to the importance of evil not only in Kant's moral theory but also as the reality that prompts us to build communities, Part VII concludes with Robert Gressis elucidating how, on Kant's view, a good person can sometimes do evil. Gressis first explains Kant's distinction between a “disposition”, an agent's “supreme maxim” that can be either good or evil, and a “propensity”, a *possible* inclination to behave in a certain manner that remains “inert” until it becomes “lively” by a person actually experiencing or indulging in that kind of behavior. After distinguishing between passive and active ways of understanding what it means to experience an evil act, Gressis examines Kant's account of how the disposition functions the first time one acts on an inclination, before it becomes lively. If the evil disposition arises through, and as a result of, the first evil act a person experiences, Kant's rigorism can be used as the basis for a plausible typology of good and evil types of persons.

Part VIII concludes Book Two with essays relating Kantian theology or philosophy of religion to another religious thinker. Continuing the theme of the previous two essays, Hans Feger relates Kant's doctrine of radical evil to his theory of freedom, focusing on Schelling's development of both theories. In Schelling's work, a Kant-inspired idea of freedom in its relation to the roles of evil and good becomes the basis for a theologically constructed system that goes beyond “religion within the boundaries of mere reason”. Of particular interest to Feger is both philosophers' attention to freedom as a noumenal and hence unprovable act that serves as an unknowable beginning of empirical acts. The very incontestability of Kant's theory of noumenal causality, though often criticized, constitutes its strength, a strength Schelling develops more explicitly than Kant himself.

Mohammad Raayat Jahromi sketches background factors influencing the development of Kant's Critical philosophy to highlight his revolutionary stance toward God and theology. Whereas traditional theol-

ogy and metaphysics had focused on God, employing theoretical reason in hopes of proving God's existence, Kant denied theoretical reason any access to the mysteries of historical religion, presenting practical reason as the philosopher's best tool for assessing the proper role of religion in human life. Jahromi compares Kant's position on a range of religious issues with that of St. Thomas Aquinas, concluding that, whereas Aquinas (and the traditions of metaphysics that followed him) promoted "theological morality", Kant developed a new, "moral theology". A review of the moral theology presented in *Religion* illustrates how Kant "opposes theoretical rationalism in the realm of historical faith on the one hand, and defends practical rationalism in the realm of moral faith on the other."

Chan-Goo Park then compares Kant and Wittgenstein on a range of issues, beginning with their views of scientific knowledge and how it relates to our ability to know the "I" that accompanies all our perceptions. Like Kant, Wittgenstein recognized that this "I" either consists of empirical contents that are not essential to the "true" self, or else refers to the transcendental limit that is never in itself knowable. As a foretaste of the essays in Book Three that examine the Kant-Buddhism relationship in greater depth, Park tantalizingly identifies this insight with what Buddhists call "emptying" the contents of the mind to obtain awareness of absolute self-knowledge. Both Kant and Wittgenstein affirm that the self is knowable only in a practical sense, as we work out our self-understanding in ethical situations. For both, belief in God does not function as an item of theoretical knowledge, but arises out of the proper consideration of our practice: either as a "postulate of practical reason" or as a "form of life".

Next, Kiyoshi Himi presents a summary interpretation of Albert Schweitzer's doctoral dissertation on Kant's philosophy of religion. Schweitzer devoted a chapter to each *Critique*, but curiously placed his discussion of *Religion* between the chapters on *CPrR* and *CJ*, rather than preserving their chronological order. Schweitzer claimed the Canon chapter of *CPR* summarizes Kant's philosophy of religion, but *CPrR* deviates from this original plan. Discussing *Religion* immediately after *CPrR*, Himi argues, enabled Schweitzer to highlight the contrast between the way these two works present the relation between morality and religion. Schweitzer shows how *CJ*'s concluding discussion of moral teleology and teleological theology laid an important foundation for the new (and originally promised) approach in *Religion*. Himi sees the key contribution of *Religion* as replacing the highly individualistic approach

of Kant's moral philosophy with a new focus on the solidarity of religious communities. Schweitzer's early exposure to Kant revolutionized Schweitzer's own interpretation of Christianity, motivating him to devote much of his life to humanitarian service as a medical doctor in Africa.

Brandon Love completes Part VIII and Book Two with an essay tracing Kant's influence on Paul Tillich's theology. Adopting the recent "Palmquistian" turn in the study of Kant's philosophy of religion, Love portrays the overall architectonic plan of Kant's Copernican Perspective as developing a theory of human personhood. After summarizing the basic transcendental elements in *Religion*, Love argues that Tillich's Christian theology exhibits markedly similar claims. Toward the end of his life, Tillich embraced a form of pluralism that Love takes to be unjustified; he suggests it be countered with "transcendental inclusivism", claiming Kant provides a set of necessary conditions for regarding any faith as genuine religion. While Love's claims may be more provocative than substantively demonstrated, they serve as an effective preparation for Book Three, where Kant scholars grapple with a range of issues relating to Asian philosophies, some of whom adopt positions that would appear to support Love's portrayal of Kant as a religious inclusivist, while others would seem to portray Kant more along the lines of a Tillichian pluralist.

#### 4. East-West Perspectives on Cultivating Personhood

While some of the essays in Books One and Two refer to East-West perspectives in the course of examining Kant's understanding of how personhood is to be cultivated, Book Three focuses on Asian themes arising out of Kant's revolutionary philosophy. An idea that emerges as a potentially unifying thread is that Kant's "revolution" is not so much a revolution (when considered in the historical context of world philosophy) as a reformation, a rediscovery of the classical roots of what philosophy ought to be. However, opinions differ.

Part IX opens Book Three with a set of essays on Mou Zhongshan, the philosopher widely recognized (in China) as the most influential Chinese Kant scholar. While the essays inevitably overlap in their coverage of Mou's most famous claim, that Chinese philosophy can fill the gap left by Kant's system by demonstrating how intellectual intuition is possible, the different angles adopted reveal fresh nuances in each case.

The first essay, by Pong Wen-berng, points out that for Kant the proper method of gaining knowledge of objects parallels the proper method of gaining self-knowledge. In both cases Kant distinguishes between the object/"I" that appears and the object/"I" in itself; knowledge of the latter is impossible for both. Mou accepts Allison's "three aspect" interpretation of object-knowledge as the best reading of Kant's texts, but claims various contradictions arise when one adopts this strategy to reflect on the three ways of referring to the "I". The main problem, for Mou, is that "in order to keep the premise of the three-aspect theory ..., Kant is forced to maintain the thesis of self-consciousness without self-knowledge"; Mou thinks "we had better adopt the three-I theory instead of the three-aspect theory, for in this way we can extend self-consciousness to self-knowledge." Pang stops short of elaborating Mou's defense of this claim, because unlike some of the following contributors, Pang regards it as untenable.

The next essay, by Mihaela C. Fistic, attempts a reconciliation of Mou's stance on intellectual intuition with Kant's. After briefly introducing Mou's account and his reasons for rejecting Kant's position (points developed more fully by Pang), Fistic delves into *CJ*'s distinction between intuitive (archetypal) and discursive (ectypal) understanding. She argues that Kant himself devotes much of *CJ* to arguing that human beings, in order to find purposes in the world and in our own lives, must interpret our experience "as if" we did have the kind of divine, intellectual intuition that, taken literally and in all its fullness, applies only to God. In this respect Kant might have been more sympathetic with Mou's attempt to go beyond Kant than Mou himself recognized.

Sammy Xia-ling Xie offers a wide-ranging assessment of Kant's view of personhood in relation to themes in Chinese philosophy. Kant's "duality of human beings" distinguishes consciousness as noumenal ("thing in itself") from our physical, sensible, and social lives as phenomenal. That the unity of these is purely "intellectual" (via pure apperception) gives rise to the problems Mou highlighted. Mou's claim that Chinese philosophy provides resources that make intellectual intuition possible for us, Xie argues, is based on a misunderstanding: what *Mou* calls intellectual intuition *Kant* calls "intellectual consciousness, self-consciousness"—i. e., the spontaneity of apperception. For Kant (as noted in Kitcher's keynote essay and throughout this collection), this spontaneity has both moral and theoretical aspects; Xie, like Mou, focuses on the latter. To resolve what he regards as a problem

for Kant, whether *criminal* activity is phenomenal or noumenal, Xie suggests adding a *third* term: between and superior to the phenomenal and noumenal we should postulate “the great flux”; this, he tantalizingly concludes, is the realm of “substance”, the essence of personhood, and the proper locus of criminal responsibility.

Annie Boisclair focuses on Mou's critique of Kant's moral theory, especially the highest good. Mou employed the Tiantai Buddhist theory of the Perfect Teaching, appealing to absolute truth as an inexpressible reality that accounts for all “dharmas” (phenomena). Mou regarded Kant as an advance on Stoic and Epicurean moral philosophies; but Kant's weakness was to appeal to God to guarantee the existence of phenomena and the proper balance between virtue and happiness. The Perfect Teaching approach requires no such appeal: “the Perfect Buddha” guarantees “the Perfect Good”. Is this an improvement? The two appeals seem equally suspicious to anyone who lacks the religious belief lurking in the background. Mou highlights several differences: whereas for Mou's Kant humans cannot influence the relation between virtue and happiness (this being wholly under God's control), “all humans have the capacity to reach buddhahood”. Mou claims the relation between virtue and happiness is analytic for us, just as for Kant's God. Although Buddhism does not tell us how this is possible, Confucianism's emphasis on morality provides concrete guidance.

Part IX's concluding essay, by Chong Chaehyun, returns to the starting point by examining Mou's criticism of Kant's theory of things in themselves as a merely conceptual correlate for theoretical knowledge of phenomena. That noumena are unknowable except as regulative ideas or (for practical reason) “postulates” is “insufficient” and “defective”: it provides no means of obtaining genuine, theoretical knowledge of the moral. Mou thinks various classical Confucian theories (e.g., Mencius' “four sprouts” and Wang Yang-ming's “innate knowledge”) solve this problem by portraying the ideal Confucian sage as possessing intellectual intuition that provides genuine *theoretical* cognition of the moral. Chong challenges this claim as being every bit as idealistic as Kant's. By contrast, Cheng Chung-ying (cf. his keynote essay, above) proposes a hermeneutic method that sees Kantian themes in the ever-changing flux of our natural life. After comparing Cheng's rich but complex approach with Mou's idealistic but unsubstantiated claims, Chong questions whether either alternative is superior to a humble affirmation of Kant's limitation on human knowledge.

Part X looks beyond Mou, to conceptions of personhood in various Chinese traditions. Noting that in Asian societies influenced by Confucianism sincerity is far from being the quaint sentiment it has become in the West, A. T. Nuyen claims “the Confucian sincere will is equivalent to the Kantian good will.” Many accounts of Confucianism ignore sincerity, because it does not feature significantly in the practically-oriented *Analects*. Yet the *Doctrine of the Mean* gives sincerity “a metaphysical significance, while the account in the *Great Learning* has epistemological overtones.” These two Confucian classics provide “an account of sincerity in terms of the sincere will that has to be cultivated in order to return to, or to be in harmony with, one’s true self, understood as the moral self.” For Confucius, “the sincere will is what brings the self in the world back into harmony with the *dao*.” Thus, the *Analects* warn against viewing the five virtues as good in themselves, apart from sincerity: “To act sincerely in the Confucian sense is to act freely and autonomously in the Kantian sense.”

Scott Stroud assesses the role of desire and the inclinations in moral cultivation. Kant is often interpreted as an enemy of inclination, promoting an idea of moral action that aims to extirpate all human desires. Stroud reviews numerous texts to demonstrate that this was not Kant’s position. Rather, Kant saw human inclinations as inherently good but naturally subject to misuse. Moral cultivation means rationally ordering our desires so they serve our moral nature rather than thwarting it. Likewise, Xunzi is often interpreted as claiming humans are evil by nature because of our natural inclinations. Stroud again cites relevant texts showing that Xunzi actually argues only that *unregulated* desire leads us astray. Xunzi thus views human nature as morally neutral. Stroud contrasts Xunzi’s intrapersonal and interpersonal reasons for concluding that unregulated desire is “unsustainable”. Without claiming their positions are identical, Stroud puts forward three theses that both philosophers embrace: inclinations challenge moral cultivation by being *disorderly*; moral education should therefore direct us away from “self-focused orientations”; and this requires both “individual initiative” and changing the “external environment”.

Mario Wenning highlights Kant’s close affinity with Daoism by comparing their respective views on nothingness, a concept that tends to be a source of insight and paradox in the East, but either ignored or concretized in the West. Wenning illustrates these tendencies by defending

three, interconnected theses: (1) it is in the nature of human reason to search for the unconditioned; (2) this unconditioned cannot be conceived of in dogmatic metaphysical terms as something existing, but is a hypothetical non-entity; and (3) the insight into the illusory nature of claims to knowledge concerning the unconditioned does not lead to epistemic despair but harbors ethical consequences.

In defending each thesis Wenning argues that “in both Daoism and transcendental idealism we witness a parallel transition from a metaphysics of non-existence to an ethics of regulative principles.” He suggests focusing dialogue on one key difference: Kant continues to appeal, even in his ethics, to “a metaphysical fact of reason . . . , whereas Daoism sees such transcendent facts as mere illusions that lead us astray rather than keeping us searching the way.”

The last two essays in Part X offer opposing assessments of Kant's relation to Chinese Buddhism. According to David Cummiskey's reconstruction of the *Lotus Sutra's* central doctrine of “skillful means”, Tiantai Buddhists face the problem of how to justify their prohibition against killing. For Kantian moral theorists, only a deontological approach can justify such constraints. Buddhist moral theory seems similar to Aristotelian virtue ethics, but its doctrine of “no self”, arising out of the doctrine of the interdependent origination of all things and the doctrine of emptiness (whereby things just are the sum total of their relations to other things), renders it radically different. Buddhists claim we can change our emotions by changing our beliefs and that uncontrolled “afflictive emotions” can negatively influence our cognitive abilities. Cummiskey shows how Buddhist arguments justifying the constraint against harming others are thoroughly consequentialist, not Kantian. Because the doctrines of Tiantai Buddhism, especially that of the “skillful means”, reveal Buddhist moral philosophy to be “a form of consequentialist virtue ethics”, one cannot be both Kantian and Buddhist.

Ellen Zhang's essay concludes Part X by comparing Kant's twofold conception of personhood (self-consciousness lets me “know I am a self, even though I do not know exactly what the self is”) with Nagarjuna's theory of “no-self or *anatman*”, whereby the self is known only through its “five components”. Since “a conception of personhood is not the sort of thing we could possibly encounter in introspection, . . . this is a good reason for denying there could be such a thing.” Chinese Huayan Buddhists follow Nagarjuna, claiming (like Kant) that “it would be logically impossible for the ‘I’ that is the subject of experience to be at the



same time its own object.” Their moral philosophy is deontological, with the “no self” replacing Kant’s (unknowable) autonomous self. For self-knowledge they turn to “a pure experience without reasoning, or an unprojected consciousness”, positing an “epistemological transformation—a new modality of knowing and experiencing” that includes “a Kant-like ‘transcendental turn’”, whereby the noumenal (“emptiness”) interpenetrates the phenomenal in an all-embracing mystical experience of oneness with all sentient beings.

Part XI relates non-Chinese Asian traditions to Kant, beginning with three essays on Indian philosophy. Emer O’Hagan argues “that Kant’s account of self-knowledge ... is psychologically implausible, and that Buddhism offers a richer account more likely to achieve its end.” After highlighting parallels between Buddha’s and Kant’s moral psychology, O’Hagan claims the former “offers a more comprehensive understanding of self-knowledge”, emphasizing “the awareness of suffering”. Kant views self-knowledge as the foremost duty to oneself; but “his inflated suspicion of self-conceit” makes him skeptical about our ability to succeed. Kant’s antidote, holding our decisions up to the light of duty, fails if self-deceit is as pervasive as he claims it is, as does his appeal to the sincerity of conscience (cf. Nuyen’s essay). For Buddha, by contrast, “appropriately framed introspective awareness can provide the connection between self-knowledge and wisdom that Kant assumes.” Buddha’s doctrine of “*anatta*” does not make an ontological commitment to a “No-self”, denying self-knowledge; rather, it promotes a disciplined technique of perceiving “Not-self” to *increase* self-knowledge and reduce suffering. Attention to suffering is more “phenomenologically trustworthy” for conveying self-knowledge than is Kantian reflection on duty.

Soraj Hongladarom examines various problems arising out of Kant’s distinction between the transcendent self (the unknowable self implied by the unity of apperception) and the empirical self (knowable through conscious experience). The Indian Buddhist philosopher, Vasubandhu (an idealist not unlike Kant) relates the distinction between the transcendent and empirical selves to the difference between grammatical subject and predicate, pointing out that the subject-object distinction is possible only for the empirical self. The problem Vasubandhu’s position poses for Kant’s is that “if we accept the argument that a necessary condition for effability is that it falls under the subject-object distinction, then the transcendental unity of self-consciousness here does not appear to be effable.” Hongladarom concludes that we cannot properly say each person

has a distinct transcendent self, nor that the same self belongs jointly to all persons, since the very conceptual schemes that enable us to make such distinctions already presuppose the presence of this self. We can say only that a transcendent self is ineffable, yet makes possible whatever is effable.

Ruchira Majumdar correlates Kant's moral philosophy to that of the Hindu holy book, *Srimad-Bhagavad-Gita*. Both encourage an approach to moral life guided by reason, focused on performing duties without consideration of the desired consequences (since duties are worthy of performing for their own sake alone), and grounded in the absolute value of the good will as the pure source of what is universally obligatory. Likewise, both uphold general principles of maxim-formation (e. g., respect for the moral law and respect for persons as ends in themselves), maintain that the virtuous person will eventually attain happiness in spite of evidence to the contrary that we find when we look only at earthly life, and defend a rational faith in the existence of God and an immortal life that permits us to understand how the ideals of morality can be realized by human beings. Most importantly (for the concerns of the present collection), both Kant and the *Gita* portray these eternal ideals as the proper path human beings must follow in order to realize their personhood.

Takayuki Kisaka begins by asserting that Kant grounds his moral theory in assumptions that are relative to his culturally-determined, Christian worldview. Even in defending his universalizability principle, Kant appeals to nature in a manner that shows his approach to be open to influence from teleological factors. To explain the relation between the formal aspects of Kant's theory and his own cultural presuppositions, Kisaka proposes a "principle of cultural value embodiment". Just as Kant was open to dialogue with Spinoza's position, despite their markedly different cultural backgrounds, *provided* Spinoza's defender embraces the universal requirements of pure reason established by transcendental idealism, Kisaka argues that Kant's appeal to "the feeling of the right and left hands" demonstrates his consciousness of the need to be open to cultural diversity in general. Kisaka concludes with the observation that Japanese culture does not, in general, embrace the Kantian postulates of God and immortality, yet this should not prevent "Kant in Asia" from being manifested with very different cultural presuppositions.

After a brief account of the role played by autonomy in Kant's moral theory, Mohsen Javadi introduces a selection of Iranian scholars who

have interpreted and assessed Kant's theory in various ways. Javadi provides an overview of six reasons typically given for rejecting Kant's theory: Iranian philosophers tend to prefer a more Aristotelian approach that takes *a posteriori* elements into consideration; they regard some legitimate moral imperatives as conditional; they call into question the very possibility of self-legislation; they deem Kant's formalistic approach insufficient for articulating moral codes; desires should not be totally discounted in considering the motivation for right action; and Kant excludes God from the internal realm of moral decision-making yet appeals to a transcendent God as the creator of reason. While Javadi admits that his overview is based on Iranian interpretations of Kant's theory—there are Western interpreters (including some in this collection) who would claim to overcome some of these objections—Western Kantians would do well to pay more attention to the incisive critiques coming out of Iran.

Part XII concludes Book Three and the entire collection with historical essays highlighting Kant's controversial relation to China. Alain-Marc Rieu emphasizes the crucial role metaphysics plays in the formation of human cultures. In his effort to ground religion in morality, Kant paradoxically “unlocked a new world; but this world closed the pre-modern world, thus refusing this new status and, in the name of morality, rejecting religion and metaphysics, the world of science, individuality, and experiment.” The essential feature of Kant's influence on world cultures, including China's, Rieu argues, is the divide between science/knowledge and morality/values. China's employment of Hegelian metaphysics to interpret its cultural norms has threatened its stability, for Hegelian dialectic requires “overcoming Confucianism as the defining character of Chinese morality and identity”. Kant's metaphysics, by contrast, is genuinely “transcultural”. Rieu claims the Kantian-European bias in the modern “morality dilemma” of self-cultivation lies on the side of knowledge.<sup>8</sup> Grounding morality in *law* enables Kant to follow a method like that of science. Because Confucianism regards morality as the ultimate ground even for knowledge itself, Rieu views Confucianism as a sorely needed supplement to Kantian philosophy, in both Asia and the West.

---

8 While this is almost undeniably true for Western culture, Kant's theory of the primacy of practical reason would appear to be an attack on precisely this feature of Western culture. See Simon Xie's essay concluding Part XII.

But how can Kant and Asia cooperate, if Kant himself had a low opinion of (for example) China, possibly even stooping to racism? The next two essays offer moderate attempts to account for this possibility, effectively counterbalancing an essay that goes to the opposite extreme, arguing that Kant was not only racist, but was himself responsible for the Eurocentrism that developed in the subsequent two centuries. Klaus-Gerd Giesen demonstrates, by means of a thorough comparison between historical accounts of the colonizing events of his day with Kant's own comments on colonialism and Asian ways of dealing with the European abuses of their hospitality, that Kant was both very knowledgeable about and deeply impressed with the Asian responses to various abusive European policies. Kant's basic argument was that, "because a world community of peoples exists, a serious violation of rights by one people or state is an attack against all other peoples." That the allegedly racist Kant (see below) held the deepest respect for Asian peoples is evidenced by his incorporation of Asian hospitality into his theory of cosmopolitan law.

Following Giesen's account of Kant's empathy with Asian perspectives on abusive European colonialism, Rein Vos scrutinizes the context of the apparently harsh critique of Chinese philosophy (and especially Confucius) in Kant's lecture notes on physical geography. Kant appears to make the shockingly inaccurate claim that no moral philosophy has ever existed in China. Vos demonstrates, however, that Kant's comments actually refer to Confucius' *political* philosophy; they are not a general assessment of his moral theory. Wolff and others had recently praised Confucius and upheld China as a shining example for Europe to follow. Yet Confucius' appeal to the wisdom of the sage kings appears to rule out Kant's Enlightenment view that public officials should think for themselves. Vos admits Kant may have misunderstood Confucius, but defends Kant's shocking comments as making sense *in context*: Kant thought Confucius honored the "political moralist"; this sharply opposes the "moral politician" required by Kant's political philosophy, with its anti-authoritarian emphasis on maximizing individual freedom through the rule of law. Does Kant's denial that any *inherently moral* idea can arise from a *political* philosopher who had called for a return to the authorities of old make him a racist? Probably not. Yet the next essay advances evidence that causes even the most charitable interpreter of Kant to pause for thought.

Peter K. J. Park's detailed study of post-Kantian approaches to the history of philosophy reveals that the first histories of philosophy to

focus on Greece as the birthplace of philosophy, to the exclusion of influence from other earlier cultures, were written by scholars who were explicitly constructing histories of philosophy along Kantian lines. After reviewing the work of one influential Kantian historian, Wilhelm Tennemann, Park examines the history of philosophy proposed by Kant's contemporary, Christoph Meiners. Though rivals, Meiners and Kant shared many ideas about race and worked as a "tag team" to introduce concepts of "race" into discussions of anthropology and the history of philosophy. While Meiners went into much more depth than Kant did, Park shows that Kant was not merely following Meiners, but was complicit in promoting racist ideas that led to "Eurocentrism"; he even influenced Meiners in at least one way, by first "using skin color as the prime marker of race". Park notes the inconsistency between such anthropological claims and the fact that Kant's Critical philosophy promotes universal ideas that supposedly apply equally to all human beings, but offers no suggestions for resolving this paradox.

Taking a step back from the problem of Kant's relation to Asia *per se*, Soo Bae Kim's essay (in German) examines the general question of how Kant's philosophy of history can be consistent with his moral philosophy, where pure morality (being noumenal) obviously cannot develop in time. Focusing on Kant's "Idea for a Universal History", Kim argues that in his anthropological and historical writings Kant views autonomy not as a given reality so much as a specifically human "task": determining the empirical conditions for making autonomy real. The philosophy of history cannot guarantee that human cultures will progress, for history as such is the account of the chaos created by human beings exercising freedom. Philosophers should not impose order onto events where there is none, but seek to cooperate with nature, in its paradoxical relationship with human freedom, by providing concrete steps human societies can follow to make the (ahistorical) ideal of the moral society humanity's final purpose. We create this idea and impose it onto history's events for the sake of our own moral self-cultivation.

Simon Shengjian Xie's concluding essay throws Park's paradox into relief by questioning whether Kant is actually a Western philosopher. Xie divides philosophy into four branches: metaphysics and ethics are typically Eastern; epistemology and logic are typically Western. Western philosophy is essentially empiricist and analytic; its metaphysics and ethics employ logical reasoning to defend claims devoid of wisdom or insight. Against those who think "philosophy" means "love of *reasoning*",

Xie echoes a key theme in the Kant-China dialogue (cf. Part IX): “True philosophy must involve intellectual intuition, not just logical reasoning and analysis. Without the former, philosophy will become shallow, purposeless, and soulless.” Like Kant, Chinese philosophy focuses on metaphysics (Daoism) and ethics (Confucianism). It “starts from the point where Kant’s philosophy ends” by providing “intellectual intuition”. Daoism, for example, acknowledges ways of understanding the Dao without knowing what it is. Xie praises Kant for displaying what Chinese philosophers *mean* by intellectual intuition (i. e., free use of imagination through non-logical methods such as analogy), even though he rejected intellectual intuition for the (Western) reason that it cannot provide knowledge.

Partly in response to the insight gained by the experiment described in essay 64, serving as the Epilogue (see §1, above), this publication refines the conference theme to reflect the actual emphasis many essays have on *cultivating* personhood, given that establishing anything like the *unity* of personhood (not only for us as individuals, but all the more so for societies or for humanity as a whole) is a task to be fulfilled (see essay 62) more than a reality to be described. That is, because of the great diversity among peoples and the resulting “crookedness” we find in the “wood” of humanity, we must constantly nurture and work at realizing the ideals of unity in diversity that both Kant and Asian philosophies show us in so many ways. Many of those who attended the Kant in Asia conference became convinced that the future of Kantian philosophy and the future of Asian philosophies must dovetail if either is to survive the pressures of twenty-first century globalization. This heretofore rarely voiced claim enters the mainstream of both Kant scholarship and East-West comparative philosophical studies with the present publication.

## *Keynote Essay to Book One*

### Kant's Spontaneous Thinker and (More) Spontaneous Agent

*Patricia Kitcher*

#### 1. Introduction

In the Preface to the *Groundwork of the Metaphysics of Morals* Kant presents speculative or theoretical reason, the faculty involved in cognition, and practical reason, the faculty involved in action, as one and the same:

A critique of practical reason, if it is to be complete, requires, on my view, that we should be able at the same time to show the unity of practical and theoretical reason in a common principle, since in the end there can only be one and the same reason (4:391).

Several years ago Onora O'Neill argued that since the categorical imperative is the principle of pure practical reason, it must also be the highest principle of pure theoretical reason.<sup>1</sup> Kant's view of the unity of reason has inspired book projects, such as Susan Neiman's, *Unity of Reason: Rereading Kant*, and Dieter Henrich's *Unity of Reason*. It also stands behind the idea of some scholars and sympathizers that Kant assimilates the spontaneity of thought and action. Henry Allison, for example, has argued for some years that Kant's notion of absolute spontaneity in action can be made clearer and more palatable by assimilating it to the spontaneity he attributes to judgment. More recently, Akeel Bilgrami has tried to clarify the role of agency in thought by assimilating it to the case of moral agency. He takes Kant to be a prime example of a philosopher who took freedom to be as essential to thought as to action.<sup>2</sup>

---

1 Onora O'Neill, *Constructions of Reason* (New York: Cambridge University Press, 1989).

2 Akeel Bilgrami, *Self-Knowledge and Resentment* (Cambridge, Ma.: Harvard University Press, 2006), 170.

By contrast I argue that we lose what is most persuasive in Kant's cognitive theory and in his moral theory by closely aligning the two. I begin by laying out a new way of understanding the cognitive theory and its relation to the "I-think" that I have developed at length elsewhere.<sup>3</sup> Then I take up the solid textual support that Allison and Bilgrami provide for their assimilations of the theories. I suggest that reading the cognitive theory in terms of the absolute spontaneity of the ethics undermines the central claims of the former—including the intimate relation between "I" and "think". I then explain how the mature ethical theory goes beyond—and must go beyond—the theory of speculative reason. Although I criticize Bilgrami's and Allison's appeals to Kant, the fault is not theirs, but his. After struggling to explain morality in terms of pure practical reason alone, he finally accepts that he has no choice but to change his position.

## 2. Judgment and the I-think

Kant's theory of cognition in the *Critique of Pure Reason* (hereafter *CPR*) is closely tied to his doctrine of the "I-think". The indissoluble expression "I-think" captures this fact: it makes no sense to talk about an "I" apart from acts of thinking and no sense to talk about such acts apart from an "I". Any plausible account of his theory of cognition or judgment or thinking must explain the intimate relation between "I" and "thinking" or "judging".

As did his predecessors, Kant presents various levels of cognition, but the cognition that is relevant to *CPR*—what he calls cognition in the proper sense of the term (A78/B103)—is conceptual cognition. It is widely agreed that the transcendental deduction (hereafter TD) proceeds by examining the necessary conditions for the possibility of experience or the possibility of empirical cognition. The TD is supposed to show that the use of *a priori* categorial concepts (and their associated principles) is a necessary condition for empirical cognition. More precisely, it is to show that employing the categories is a necessary condition for the cognition distinctive of humans—conceptual cognition.

An early essay where Kant criticizes the logician G. F. Meier's suggestion that animals use concepts provides insight into how he understands concept use.

---

3 Patricia Kitcher, *Kant's Thinker* (Oxford: Oxford University Press, 2011).



[Meier's] argument runs like this: an ox's representation of its stall includes the clear representation of its characteristic mark of having a door; therefore, the ox has a distinct concept of its stall. It is easy to prevent the confusion here. The distinctness of a concept does not consist in the fact that that which is a characteristic mark of the thing is clearly represented, but rather in the fact that it is recognized as a characteristic of the thing. The door is something which does, it is true, belong to the stall and can serve as a characteristic mark of it. But only the being who forms the judgment: **this door belongs to this stable** has a distinct concept of the building, and that is certainly beyond the powers of animals.

I would go still further and say: it is one thing to **differentiate** things from each other, and quite another thing **to recognize** the difference between them (2:59–60).

Although this essay is early, Kant's Logic Lectures offer a similar picture in characterizing the relation among "discursive" cognition (A68/B93), concepts, and their characteristic "marks":

From the side of the understanding, human cognition is discursive, i. e., it takes place through representations which take as the ground of cognition that which is common to many things, hence through marks as such. Thus we cognize things **through** marks and that is called **cognizing**, which comes from being acquainted.

A mark is that in a thing which constitutes a part of the cognition of it, or—what is the same — a partial representation, *insofar as it is considered as a ground of cognition of the whole representation*. All our concepts are marks, accordingly, and all thought is nothing other than a representing through marks (9:58, my italics).

The text Kant used for the course describes one use of marks (16:296). Marks are a ground of cognition, because they enable cognizers to differentiate the object of cognition from other things (24:113; 16:299). For example, it is orange, other things are not. By contrast, Kant maintains that marks not only have this "external" use but also an "internal use". In the latter case marks or partial representations are the ground for applying the whole concept to the object, not via identity and differences with other objects, but via identity of the marks (16:298 [R2282]). The internal use of marks is not a matter of differentiation but of derivation (16:299). For example, the concept "body" might include the marks "impenetrable", "extension" and so forth (A106) and so be applied through the tacit judgments "this thing is impenetrable", etc.

For a representation to be a mark, it must be considered as such. It is not that cognizers must have the concept "mark" or the concept "concept" or the very abstract concept "representation". Rather, they must recognize that a mark—say "impenetrable"—is a partial ground or basis

for their application of the concept “body”. It is part of why they call something a “body” or part of what they presuppose in calling something a “body”. Even when not considered in relation to a complex concept, but just on its own, a mark is still a basis of cognition, because it is the ground of cognition of the objects in its extension (9:96):

As one says of a **ground** in general that it contains the **consequences** under itself, so can one also say of the concept that as **ground of cognition** it contains all those things under itself from which it has been abstracted, e. g., the concept of metal contains under itself gold, silver, copper, etc.

“Metal”, for example, is a ground of cognition of copper things, because it classes them together with other metals. In describing a copper kettle as “metallic” a cognizer implies that it is similar to some things and different from others. To recognize a mark as such is to use the term with that understanding.

Both animal and human sensory systems detect similarities in the properties of objects. Those who are capable of discursive cognition recognize their representations of, e. g., the color orange as presenting a humanly detectable property common to many things and (so) as marks—as things that can be offered in answer to the question of why they group those various things together (and exclude others) and of why they label something as a particular kind of fruit, an “orange”.

With a better sense of Kant's understanding of concept use, we can turn to the relation between concept use or judging and the “I-think”. It is surely no coincidence that the theory of apperception is introduced in the A Deduction in the section on “Recognition in a Concept”. This is a useful discussion, because it is one of the rare occasions where Kant illustrates his epistemological claim with an extended example. The example is counting:

Without the consciousness that what we are thinking is the same as what we thought an instant before, all reproduction in the series of representations would be futile ... the manifold of representations would never make up a whole, because it would lack that unity that only consciousness can impart to it. If, in counting, I forget that the units that now float before my mind or senses were added together by me one after another, I should never cognize the amount being produced through this successive addition of unit to unit; nor, therefore, would I cognize the number. For this number's concept consists solely in the consciousness of this unity of synthesis (A103, amended translation).

His first point is that the mere ability to reproduce the series of stroke symbols, say four of them in a row, is insufficient for cognition. He does not refer to animals here, but the contrast is useful. An ox could have an image of four stroke symbols, but he could not recognize its contents under the concept “4”.

Kant believes that concepts are associated with rules (e.g., A106). Since mathematical concepts are usually understood as having definitions, his example may suggest that the associated rules are necessary and sufficient conditions for the applicability of the concept. In fact, he thinks that neither empirical concepts nor categories can be defined (9:41, A727/B755–6). The associated rules are not definitions, but partial explications of the concept (A729/B757). He thinks of these rules as universal, but for the case of empirical concepts at least, it would not matter to his theory if the rules were probabilistic. Once the suggestion of necessary and sufficient conditions is rejected and the rules are allowed to be probabilistic, Kant’s assumption that concepts are associated with rules can be seen as a version of the contemporary view that concepts stand in inferential relations to other concepts and can be used only by individuals who explicitly or implicitly recognize those relations. The rules indicate some of the relations. Alternatively, the rules give the marks of the concept. In the case of concepts that are either not complex or not clear (where the subject does not know the inferential relations), the rule would be that of the external use of marks—the rule that the concept indicates a property that can be detected by humans and is common to this object and others.

Using the counting rule (or any rule associated with a concept) involves a number of skills. Counters must be aware of their performance in such a way as to catch possible errors and, in this case, to know where they are in the process. Kant presupposes all this in making his second and positive claim: counters need to be conscious that they designate the first stroke symbol as “1” etc. in order to cognize the amount. They need to be conscious of their representations of “1”, “2”, etc., as “marks” or “partial representations” that provide the ground or basis of his conceptual representation “4”. The last sentence of the passage notes that, in this case, being conscious of applying the counting rule to the members of a set is not only necessary for applying the concept to the set, it is also sufficient.

Kant elaborates the account in the further discussion of this example:

The very word concept could on its own lead us to this observation. For this one consciousness is what unites in one representation what is manifold, intuited little by little, and then also reproduced. Often this consciousness may be only faint, so that we do not [notice it] in the act itself, i. e., do not connect it directly with the representation's production, but [notice it] only in the act's effect. Yet, despite these differences, a consciousness must always be encountered, even if it lacks striking clarity; without this consciousness, concepts, and along with them cognition of objects, are quite impossible (A103–4).

He allows that thinkers do not have to pay much attention to individual steps, adding up the stroke symbols little by little. Still, they must be conscious of the judgment “4” as the effect of applying to the represented units the rule that a set has the size of “4” if and only if it consists in four units. The discussion concludes with the very strong claim that without this consciousness, cognition of objects would be impossible.

Given Kant's view of how concepts are employed, we can understand why he thinks conceptual cognition requires consciousness of acts of judging. Conceptual cognition requires that partial representations, “1”, etc., are not merely representations that float before the mind. They must be understood as partial representations, as the basis of the whole representation “4”. If cognizers were not conscious of the act of applying the concept “4” on the basis of the representations of the units, then they would not know the basis of their judgments and so would fail to be rational cognizers. With arithmetical concepts the rules offer necessary and sufficient conditions for the applicability of the concept. So a cognizer could infer that the number of units in the set is “4”. Where the rules are merely partial explications, e. g., “bodies are impenetrable,” the judgment that  $x$  is a body has “ $x$  is impenetrable” as its basis or partial ground, although the judgment “ $x$  is a body” is not a valid inference from that ground.

A few paragraphs later, Kant offers the second example of “body”. The concept “body” necessarily involves representations of impenetrability, shape, etc. (A106). He explains that a concept can be “a rule for intuitions only by representing, when appearances are given to us, the necessary reproduction of the manifold and hence the synthetic unity in our consciousness of these appearances” (A106). That is, the materials in the partial representations, “impenetrability” etc., must be “reproduced” in the resulting representation “body”; alternatively, the latter representation must be understood as being built up out of the partial representations that are “repeated” in it. Kant then abruptly claims

that any necessity must be grounded in a transcendental condition, so we need to find a transcendental basis for the “unity of consciousness in the synthesis of the manifold” (A106). By a “transcendental basis” he means an *a priori* representation that is required for cognition. Then he makes the dramatic pronouncement: “This original and transcendental condition is none other than transcendental apperception” (A106–7). We can make sense of the *prima facie* puzzling claim that concept use requires “transcendental” apperception by considering the rule associated with “apperception” or with the “I-think”—namely, the rule that different representations must belong to a common I-think (A117a). For rational cognizers to apply the concept “body”, they must see their judgment, “x is a body” as having as partial grounds, “x is impenetrable,” “x has shape,” and so forth. That is, I must recognize that my judgment depends on and would be impossible without the partial representations that are its grounds and that those partial representations would be impossible as grounds of cognition unless they were the basis of the judgment. Thus, in recognizing that I can and do assert “x is a body” only because I can assert, for example, “x has shape,” I also recognize that the two representations, *qua* representations, must belong together and so are instances of the “I-rule”, the rule that different representations must belong to some common subject. More generally, in any case where cognizers can engage in “cognition proper” where they know the grounds of their judgment, they must understand the representations involved as necessarily connected to each other (and so to a common subject).

Kant makes the same point in the B deduction. The cognition distinctive of humans requires that they understand combination as such (B130–1); this in turn requires the unity of apperception or self-consciousness. He also argues, *contra* Locke, that humans can recognize their identity through time *only* by being aware of their acts of combining representations (in representations of objects):

[T]he empirical consciousness that accompanies different representation is intrinsically sporadic and without any reference to the subject’s identity. Hence this reference comes about not through my merely accompanying each representation with consciousness, but through my adding one representation to another and being *conscious of their synthesis* (B133, my italics).

Kant's objection is that Lockean consciousness—the consciousness that is inseparable from thinking<sup>4</sup>—is momentary or episodic. As such, it cannot provide a basis for representing a common or enduring subject.

How does being aware of these acts of combination or synthesis enable cognizers to represent their identity? Kant's point is easier to see in concrete cases, so we can return to the counting example. Through being aware of their act of combining, counters recognize that they have applied the “whole” concept or made the judgment on the basis of applying the counting rule to “partial” representations that hover before the mind. They recognize those partial representations as the grounds of their concept application. Thus, they recognize that the conceptual representation or judgmental state would not exist without those representations. In this way, cognizers recognize the representations they are conscious of as instances of the apperception rule, the rule that different representations must belong together. Consciousness of the act of synthesis is crucial for conceptual or discursive cognition. Without it cognition of objects would be impossible, because the marks of a concept could not be regarded as the grounds of cognition. But with it cognizers are also able to recognize the unity of their consciousness.

### 3. Allison's “Taking as” Account of Kant's Theory of Judgment

As noted, Allison's tries to make Kant's incompatibilism in ethics more acceptable by assimilating it to his theory of judgment.

In order to understand Kant's seemingly gratuitous insistence on a merely intelligible moment of spontaneity in the conception of rational agency, we must look not to his moral theory or motivational psychology but rather to his views on the spontaneity of the understanding and reason in their epistemic functions.<sup>5</sup>

He frames his discussion of the “freedom” involved in judging in terms of Wilfrid Sellars' distinction between “relative” and “absolute” sponta-

---

4 John Locke, *Essay Concerning Human Understanding*, ed. P. H. Niddich (Oxford: Clarendon Press, 1690), 2:27.9.

5 Henry Allison, *Kant's Theory of Freedom* (New York: Cambridge University Press, 1990), 36.

neity.<sup>6</sup> I set up the discussion in terms of Sellars' categories, but only (later) to reveal the inadequacy of this distinction for understanding Kant's views of autonomy. Sellars defined a "relatively spontaneous" act as one that is set in motion by a "foreign cause" and that follows a (preset) routine. He clarified the notion by making an analogy with a computer that receives input and then follows an algorithm in determining an answer.<sup>7</sup> He defined "absolute spontaneity" by contrast with "relative", *viz.*, as a spontaneity that does not involve a foreign cause and a routine for dealing with material supplied by that cause.

In objecting that my previous accounts of transcendental psychology failed to deal adequately with Kant's normative concerns in the TD<sup>8</sup>—a fair enough criticism—Allison argues that a merely relative spontaneity can never be enough.

As long as cognition is viewed as essentially a matter of being in the appropriate cognitive state, a merely relative spontaneity is all that one need assume. This ceases to be the case, however, once one grants that cognition requires conceptual recognition or *taking reasons as reasons*. Since this is an act that the subject must perform for itself (self-consciously) rather than a cognitive state in which it finds itself, it follows that we must assume an absolute and not merely a relative spontaneity in order to conceive of its possibility.<sup>9</sup>

He provides an especially clear account of how he understands the spontaneity of judging in a recent essay:

The basic point is that to consider oneself as a cognizer is to assume such [absolute] spontaneity. This is because to understand or cognize something requires not simply having the correct beliefs and even having them for the correct reasons, it also involves a capacity to take these reasons (whether rightly or wrongly) as justifying the belief.<sup>10</sup>

In many ways, my current analysis agrees with Allison's. As I understand Kant's dissection of recognition in a concept or judging, cognizers see (or can see) the partial cognitions as grounds of their judgment; other-

---

6 Wilfrid Sellars, "this I or he or it (the thing) that thinks", in *Kant's Transcendental Metaphysics*, ed. J. F. Sicha (Atascadero, Ca.: Ridgeview, 1972/2002).

7 Sellars, 356.

8 Henry Allison, *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (New York: Cambridge University Press, 1996), 58.

9 Allison, *Idealism and Freedom*, 64.

10 Henry Allison, "Kant and Freedom of the Will", *Cambridge Companion to Kant and Modern Philosophy*, ed. P. Guyer (Cambridge: Cambridge University Press, 2006), 389.

wise they would not enjoy rational cognition. Conceptual cognition is not simply a matter of having representations that would instantiate a rule for bodies, e. g., “bodies are impenetrable,” but of grasping the logical relations between the concept that is used, “body”, and the partial representation, “impenetrability”.

To appreciate the differences between our views, it is helpful to recall how the B deduction sets up the problem to be solved by the theory of apperception or “I-think”. We can understand the possibility of rational or conceptual cognition only if we can understand combination as such (B130–1). On my analysis, the argument is fairly simple. To recognize a combined representation as such a cognizer must grasp it as having been combined from other representations. Cognizers are able to do that in the case of judging, because they can be conscious of combining the partial representations in the complex representation, thereby recognizing the representations as necessarily connected and thus as instantiating the “I-rule”.

As I understand it, Allison’s “taking-as” construal of Kant’s theory of judgment or recognition in a concept would not solve the problem of understanding combination as such. It would make it insoluble. Absolutely spontaneous cognizers would see the relation between their representations of “impenetrability”, etc., and their representation of “body” as contingent. They could judge that “x is a body”—or not—regardless of whether they could judge that “x is impenetrable.” They would not see various representational states as necessarily connected and so as belonging to a unity of apperception. On this account, there is no understanding of combination as such, and no role for the unity of apperception as a necessary condition for combination understood as such.

Allison’s idea seems to be that rules connected to concepts are analogous to maxims of action in functioning as norms for cognizers, because they freely choose to apply them—or not. Of course, cognizers can refuse to make judgments by recognizing the rules they associate with concepts and choosing to reject the rule rather than to make or withhold the judgment. Given this option, they could see their judgments, e. g., “x is a body,” both as based on reasons, when they make them, and as free, because they are free to reject the rules for “body”. Although this is true and important for the kind of self-criticism and self-improvement that is central to Kant’s view of cognition, it cannot be what Allison has in mind. Either cognizers would judge in accord with a rule and so not be absolutely spontaneous, or they would reject



the rule for some reason, most likely, because it is no longer seen as the best means of distinguishing groups of things (A728/B756)—and still not be absolutely spontaneous. In his *Logic Lectures* Kant rejects the Cartesian view that cognizers can withhold belief at will: “The will does not have any influence immediately on holding-to-be-true; this would be quite absurd” (9:73; 16:398). In the case of judging or “objective” holding-to-be-true (A822/B850), the absurdity would be twofold. The “judgment” would not be an instance of rational cognition, because such “thinkers” would not know the reason grounding the judgment: they would not understand their states as necessarily connected to each other and so would not understand themselves as thinkers.

#### 4. Evidence for Assimilating Thought and Agency

Although Allison’s assimilation of Kantian judging to Kantian free action leads, I believe, to a distorted view of the former, it is well-grounded in the texts. I cannot present all of that evidence, but only a key passage that will show where the error lies—namely with Kant. Other passages that Allison cites can be handled in a similar fashion.

Bilgrami’s aim is not Kant interpretation, but the development of what he takes to be a Kant friendly view. He appeals to two passages, one from *CPrR* and one from *GMM*. The *CPrR* passage offers Kant’s slightly defensive account of the relation between the theories of *CPR* and *CPrR*.

Accordingly, considerations of this kind which are once more directed to the concept of freedom in the practical use of pure reason, must not be regarded as an interpolation serving only to fill up the gaps in the critical system of speculative reason (*for this is for its own purposes complete*) ... This remark applies especially to the concept of freedom, respecting which one cannot but observe with surprise that so many boast of being able to understand it quite well and to explain its possibility, while they regard it only psychologically, whereas if they had studied it in a transcendental point of view, they must have recognized that it is not only **indispensable** as a problematic concept, in the complete use of speculative reason [as well as completely **incomprehensible**.] (5:7)<sup>11</sup>

This passage seems to me to assert the opposite of Bilgrami’s thesis. In the text I italicize, Kant denies that “freedom” is needed to fill gaps in

---

11 Cited in Bilgrami, 170–1. I add the part in square brackets from the original text; my italics.

his critical system of speculative reason. Presumably this is because he believes he has accomplished his *CPR* goals of explaining the possibility of empirical cognition and of diagnosing metaphysical error without a defense of freedom. Why then is the concept of freedom indispensable in the complete use of speculative reason? A transcendental perspective looks at the sources of cognition. It reveals that reason looks for ever higher causes, a regress that could be halted only with the discovery of a cause that is not caused, but is an original beginning. The concept of freedom is indispensable in the literal sense that it is needed to characterize the strivings of reason that are the source of metaphysical error. What *CPR* left open was not the possibility of a further necessary condition for empirical cognition, but a defense of the idea of freedom as something other than a necessary element in the description of error. It did not answer the gaping question of why reason not only encourages humans to look for ever deeper explanations, but pushes them to look beyond the bounds of empirical cognition.

The passage Bilgrami cites from *GMM* is the crucial discussion of “thinking and acting under the idea of freedom” and is much more helpful in supporting his assimilation of Kant’s theory of thought to his theory of action. Allison also appeals to this important material to support his claims about absolute spontaneity in cognition.<sup>12</sup>

1. Now, one cannot possibly think of a reason that would consciously receive direction from any other quarter in regard to its judgments, since the subject would then attribute the determination of his judgment not to his reason but to impulse. 2. Reason must regard itself as the author of its principles independently of alien influences (4:448, my numbering).

The portion I label “1” explains that it is impossible to understand thought (or reasoning) that is directed from the outside as thought (or reasoning). If an alien thought is introduced (perhaps a suggestion from a colleague), then that thought could be evaluated by one’s understanding and/or reason and adopted or rejected. That would still be thinking. What Kant regards as impossible as thinking is being conscious of an impulse and then conscious that that impulse has produced another mental state, say, a judgment. We can see why Kant holds this view from the *CPR* materials already presented. It is part of his theory that judging requires consciousness of synthesis so that the product of thought is understood as being produced from partial cognitions in ac-

---

12 Allison, *Idealism and Freedom*, 128.

cord with rules. Bilgrami is right that Kant dismisses passive watching of mental states as thinking, because the resulting “thought” has no “I” attached to it. And, it has no I attached, because of the passivity, because there was no self-conscious act of thinking. But none of this implies that thinking is absolutely spontaneous—as acting must be for Kant.

Let us now turn to the piece of text marked “2”. What does Kant mean in describing reason as the author [*Urheber*] of its principles? One suggestion is that reason either creates or chooses its principles. On this reading, Kant’s position would be that reason can be (practical) reason only if it can regard itself as the creator/chooser of its own principles. Since the passage is part of a general discussion of judgment, Allison and Bilgrami attribute the same position to him for the case of speculative reason. Given the text (and others), their attributions are reasonable. Nevertheless, the assimilation should be resisted for two reasons, one that comes from the wider context of his ethical theory, the other from the wider context of his views about rational autonomy.

It cannot be Kant’s considered view that reason chooses its own principles, and so is absolutely spontaneous, even for the case of practical reason. As we have seen, interpreting Kant as a doxastic voluntarist severs the link between rational or conceptual cognition and the unity of apperception that the TD endeavors to forge. Reading him as a voluntarist about the principles of practical reason is even more damaging to the central claims of his ethics. Without the assumption that the same moral law lies *a priori* in the practical reason of all rational agents, nothing in his ethics makes sense. I give two examples.

Consider, first, the development of Kant’s ethical theory. It is widely accepted that he changed his views about freedom between *GMM* and *CPrR*. He gives up the argument about the impossibility of acting (or thinking) except under the idea of freedom, the argument offered in the passage we have been considering. The new argument moves from agents’ consciousness of the moral law to their recognition of their freedom to act on it (5:30). More importantly, having realized that practical reason could produce only morally good actions, he introduces a faculty of choice to leave open the possibility that an agent could do otherwise in exactly the same circumstances (5:100). In *Religion within the Limits of Mere Reason* he characterizes this faculty more fully as one that decides between the dictates of reason’s moral law and the demands of self-love (6:18–25). Had Kant understood reason as operating in voluntaristic mode, then the latter change would be unnecessary. In exactly the same circumstances a rational agent might apply the moral law to the

maxim or not. It is only the relentless operation of practical reason in applying the moral law that forces the addition of a new faculty to preserve the possibility of doing otherwise and, so, of absolute moral responsibility.

Second, attributing a voluntaristic faculty of reason to Kant makes it impossible to understand the final form of his moral theory captured in the great innovation of *CPrR*, the “fact of reason”. This doctrine has been a source of puzzlement, and even derision, but recent analyses explain how it presents a clear and compelling account of moral phenomenology. In “The Concept of Moral Insight and Kant’s Doctrine of the Fact of Reason” Dieter Henrich lays out the complex picture behind the doctrine.<sup>13</sup> Kant’s view is not simply that, when confronted with morally parlous situations, agents are aware of the moral law. Rather, they are aware that, for example, their proposed action is required by the moral law, an awareness that is experienced as the demand of universal rationality for the good, and simultaneously produces the appropriate motive of respect.<sup>14</sup> Marcus Willaschek agrees with Henrich on the complex structure of the fact of reason and presents a clearer picture of the essential precursor to Kantian moral decision. The consciousness of the moral law captured in the “fact of reason” is a consciousness of the demanding character of the law.<sup>15</sup> Agents do not have a merely intellectual appreciation of the conflict between their possible actions and the moral law. Agents are also conscious of their grasp of the relation between the law and their proposed action as a force moving their will toward or away from the action. Had Kant believed that agents employed the principles of practical reason freely—that they could choose to incorporate and use its principles in their moral deliberations or not—then he would have had to forego a central teaching of his ethics. He would have had to give up the idea that all moral agents are *ipso facto* conscious of the moral “ought”.

We can sort out the disparate claims of the *GMM* passage if we look back at them from the vantage point of the ethical theory in the *Religion* book. According to the later work, in addition to the moral law that is

---

13 Dieter Henrich, “The Concept of Moral Insight and Kant’s Doctrine of the Fact of Reason”, reprinted in *The Unity of Reason* (Cambridge, Ma.: Harvard University Press, 1960/1994).

14 Henrich, 73, 83–5.

15 Marcus Willaschek, *Praktische Vernunft: Handlungstheorie und Moralbegründung bei Kant* (Stuttgart and Weimar: Metzler, 1992), 188.

given within every rational agent (6:27n), human beings have a disposition (*Gesinnung*) to incorporate the moral law as their maxim of action or not (6:24). Kant goes on to explain that even though the disposition can be thought of as innate, that “does not mean that the disposition has not been earned by the human being who harbors it, i. e., that he is not its author (*Urheber*)” (6:25). At this point, he thinks he has everything he needs to explain the possibility of genuinely free moral action. When people follow the moral law, they are following a principle of their own reason and, so, are not compelled or even influenced by alien sensory factors. Their ultimate responsibility for an action arises not because they are the creator of their own reason—as if different human beings could have different rational faculties. Rather, they are responsible because they are the author of their disposition to incorporate the judgment of their reason in their action. As Sellars notes, these are two very different senses of freedom: “The concept of autonomy must not be confused with that of free choice”.<sup>16</sup> He thinks that both are crucial to Kant’s position—but that autonomy represents the deeper sense of freedom for him.<sup>17</sup>

The *GMM* passage is difficult to interpret, not because Kant confuses autonomy with free choice, but because he tries to explain both in terms of a single faculty: Reason operates independently of alien influences and it is the creator or chooser of its own principles. In trying to do everything with reason, he gives the misleading impression that he thinks that reason freely creates or chooses its principles. But that is not the considered view of this ethical and epistemological Rationalist. It is a deep part of his ethics that agents are not free to turn off, or even to turn down, the voice of reason.

It is also an important part of Kant’s mature ethical theory that rationality is not sufficient for moral personality. Here is the canonical formulation of this point in the *Religion* book (6:26n):

For from the fact that a being has reason does not at all follow that, simply by virtue of representing its maxims as suited to universal legislation, this reason contains a faculty of determining the power of choice unconditionally, and hence to be “practical” on its own. The most rational being ... might apply the most rational reflection to those objects (incentives) ... without thereby even suspecting the possibility of such a thing as the absolutely imperative moral law which announces to be itself an incentive, and, indeed, the highest incentive. Were this law not given to us from within,

---

16 Sellars, 361.

17 Sellars, 362.

no amount of subtle reasoning on our part would produce it or win our power of choice over to it. Yet this law is the only law that makes us conscious ... (of our freedom) and thereby of the accountability of all our actions.

That is, consciousness of freedom is a special feature of the moral law of practical reason and not something that characterizes reason more generally.

Moral agency involves two features that have no place in theoretical reason. Practical reason is not merely rational, but practical, because it contains an imperative that is experienced as moving the will. Moral agents bear ultimate moral responsibility, because they have a disposition to follow or flout the voice of reason. Although theoretical reason lacks these two requirements for absolute freedom and responsibility, it possesses a kind of autonomy that practical reason lacks—or at least a kind of autonomy that Kant could not argue that it possessed. As Karl Ameriks observes, Kant came to accept that he could not provide a deduction for the moral law. Both the moral law and the necessary assumption of freedom are “established” only through the fact of reason. For this reason, Ameriks characterizes Kant’s practical philosophy as “dogmatic”, as opposed to his “critical” epistemology.<sup>18</sup>

By contrast, theoretical reason is free of “alien” influences at four levels: Reason thinks in accord with its own principles (as opposed to principles borrowed from experience), it makes those principles explicit (by using its own principles), it evaluates some of its principles by appeal to others, and it can evaluate its most basic principles (and establish their legitimacy through a TD). Practical reason is autonomous at only the first three levels.

Given the importance of the self-critical aspect of reason to Kant’s theoretical and ethical projects, Sellars’ categories of “relative” and “absolute” spontaneity are too simple. Reason does not operate by the sort of rote following that Sellars seems to include under “relative spontaneity”. Neither, however, does it involve absolute spontaneity. A self-critical reason applies some of its principles to the evaluation of others. Kant believes that humans are autonomous in thought and action, because they possess a self-critical reason. A self-critical reason, one that uses some of its principles to criticize others is, however, entirely different from a self-created one—as Sellars well understands. His other distinc-

---

18 Karl Ameriks, *Kant's Theory of Mind*, second edition (Oxford: Clarendon Press, 2000), 218.

tion, between autonomy and free choice, is better able to capture the relation between theoretical and practical reason as Kant understands them. The freedom that Kant takes to be common and essential to speculative and moral reasoning is freedom in the sense of “autonomy”; the *differentia* of moral deliberation is that it involves judging that moves the will and “free choice” as well as the autonomy of practical reason.

Kant neither carried out his project of showing that theoretical and practical reason is one and the same nor left the task for sympathetic readers. When he realized that the moral law could not be deduced from the freedom required for rational thought, he changed his focus to the source of human awareness of freedom. He traced that awareness not to a feeling of voluntariness, but to one of obligation. In tying human freedom to practical obligation and in recognizing the limits of practical reason to deduce its principles, he demonstrated that theoretical and practical reason differ dramatically in what they can accomplish.<sup>19</sup>

---

19 Many thanks to Jens Timmerman and Kwan Tze-Wan, whose knowledgeable and insightful comments enabled me to improve the paper I gave at the conference.

## *Keynote Essay to Book Two*

# Metacritique of Practical Reason: Back from Kant's Universalized Egocentrism via Kongzi's Moral Reciprocity and Mengzi's Compassion to Huainanzi's Reciprocal Resonance and Zhuangzi's Ethos without Ego

*Günter Wohlfart*

## The West: Kant

### I.1 Kant in Asia

*Kant lebt*—Kant lives; that is the allusive title of a book that appeared in 2004, 200 years after Kant's death. The epigonic Neo-Kantians celebrate his resurrection every year. Kant is the national hero of German philosophy. The most famous and influential German philosophy professors are little Kantians, "Käntchen" as Jean Paul called them—first of all my, so to speak, doctor-stepfather Jürgen Habermas.

Kant's moral philosophy, his so-called "practical reason" appears as the politically correct philosophy in Germany. Kant is Germany's best-selling philosophical export. But Kant's tree of knowledge is not only an evergreen in Germany. Kant's reloaded moral philosophy seems to become the proto-global-morality with a claim for universal validity. Kant seems to be omnipresent.

Kant is in Asia. Kant is in China—in several up-dated versions. From my own experience during the last few years I can mention two examples: the Academia Sinica in Taipei as well as the Academy of Social Sciences in Shanghai are breeding grounds for little Neo-Kantians. Kant is in Hong Kong, too. He is here, right now. I agree with the Chinese writer Zhou Derong who already said some years ago:



The spiritual achievements of the Chinese in the last two decades are solely based on Kant's critical theory of reason. Ironically enough Kant's comeback in the early eighties had to do with Deng Xiaoping's statement, which roughly reads: practice is the sole criterion for verifying the truth (of a practical theory).<sup>1</sup>

Unlike Kant, I am deeply convinced this is true. To vary Hegel's words: the truth of a practical theory is practice. A pure practical theory that is not practicable is untrue. It is only theoretically practical—like Kant's "practical" reason. I am convinced that Kant's main principle of morality is in fact unsuitable to regulate moral action. It even might be dangerous. Why? I'll try to show that in the first half of this essay. Because I believe the "source" of Kant's critique of pure reason is so pure, so distilled that it is impossible for fish to live in it, I think Kant's critique itself must be criticized.

## I.2 Critique and Metacritique

Remember Kant's prophetic admonition in the Preface to his *Critique of Pure Reason* (Axi): "Our age is the true age of criticism, to which everything [sic! G. W.] has to submit." Yes, I agree: *everything*, even criticism itself. To follow the consequences of Kant's own critique finally means to go beyond Kant and to be consequent is, following Kant, the highest virtue of a philosopher.

You remember the famous words from the end of the *Critique of Pure Reason*: "Der kritische Weg ist allein noch offen"—the *Critical* way alone is still open". Yes, I agree, but I would like to go with Kant one step further than Kant and say: the *metacritical* way alone is still open. Especially a metacritique of practical reason is necessary.

The term "metacritique" was created by Kant's contemporary and his first criticizer, Hamann, and later on borrowed by Adorno. Metacritique means critique of critique. Without this critique of critique, criticism is in danger of becoming itself a form of dogmatism: dogmatic criticism. I repeat: the task is to go *with* Kant beyond Kant, to climb on Kant's shoulders without falling back *behind* Kant into bad old dogmatic metaphysics. And by the way, Kant himself emphasized that it is not at all unusual to understand a philosopher better than he understood himself (B371).

---

<sup>1</sup> Zhou Derong, *Frankfurter Allgemeine Zeitung*, 21/7/2004.

Especially as for moral philosophy my *credo* is, I repeat: the *metacritical* way alone is still open. We have to go this way, if we want to avoid the dogmatism of an uncritical criticism. We have to go this way, if we want to avoid the dialectics of an unenlightened enlightenment, for this, as Adorno showed in his *Dialektik der Aufklärung*, is totalitarian.

The *apex* of Kant's *Critical* theory was his critique of an uncritical, unenlightened use of *God* as the crux of dogmatic metaphysics. But what about *metacritique*? What is the crux of dogmatic criticism? What is the dogmatic core that hides under the guise of critique and enlightenment? Could it be that this crux is the idea of our *Ego*, the same idea that also lets us be a *person*, as Kant says? Could it have something to do with the supposed "unity of human *personhood*"?

### I.3 The Modern Ego as the Executor of the Christian God

The first and presumptuous sentence in Kant's *Anthropologie in pragmatischer Hinsicht* reads—I translate (7:117):

That man in his conception can have the 'I' elevates him endlessly above all other creatures on earth. This is the reason why he is a *person* and because of the unity of his consciousness despite all changes which may happen to him [he is] one and the same person, that is a living thing which is, concerning its rank and its dignity, totally different from things like animals, which are not endowed with reason and which one can treat as one likes.

Kant in the footsteps of his protagonist of modernity, Descartes. The "I" elevates human beings *endlessly* above all other living things.

The pathos of this "ascension" of the Ego reminds us that this Ego still reflects the glory of God. The modern Ego is a creation in the image of the Christian God. Our person with its supposed *unity* is the *persona*, the mask of God, the mask of the Christian God as the only true one God with his supposed *uniqueness*. God was the highest being, the transcendent *apex* of dogmatic metaphysics in pre-modern times. In modern times the Ego replaced God and inherited his diseases.

Since Descartes, the forerunner of enlightenment, the Ego is the *credendum*, the first article of faith in modernity. *Ego cogito, ergo sum*—this was Descartes' famous fallacy, something like an ontological Ego-proof.

Nietzsche's later diagnosis was: "God is dead." Maybe it was a bit overhasty, because God revived. God revived in the shape of our modern Ego with its omnipotence-fantasies. A resurrection of God in the

form of the erection of our modern Ego. I'm tempted to call this our "Egod", if you permit me this neologism.

The Ego, the I, the subject, became the *ens realissimum*, the most real being in modernity. The Ego is the executor of the Christian God. The enlightenment of the Ego is the profane, mundane metamorphosis of the "light" of God. Egocentrism is secularized theocentrism. I repeat: God was the highest being, the *transcendent apex* of dogmatic metaphysics. The Ego is the highest being, the *transcendental apex* of dogmatic criticism.

Kant calls the "synthetic unity of transcendental apperception" (B133) the "highest point, on which all use of reason, the whole logic itself and after it, the transcendental philosophy must be fixed, yes, this capacity is reason (*Verstand*) itself." What a pathos! This unity of apperception is for Kant the "vehicle of all concepts", the "pole star" of his philosophical horizon. Or we even could say that the Ego is the sun, enlightening our modern cognition.

The *Copernican turn* was that the earth revolves around the *sun* and not vice versa. The *Kantian turn*, his "Revolution der Denkungsart" was that everything revolves around our *Ego* and its cognition. Copernican heliocentrism and Kantian egocentrism.

Once upon a time we believed in *God* and its *uniqueness*. In modern times we believe in our *Ego* and its identity and *unity*. But our Ego is a black box, an *asylum ignorantiae*—like God.

What about the supposed "unity of apperception" as the reason of the unity of our person? Is the *synthetic* unity of apperception, this vanishing point where the lines of thinking seem to intersect—is this supposed *synthetic* unity in truth only a *synoptic* illusion? Is the transcendental apperception a transcendental illusion? If I say "transcendental" I use it in the strict Kantian sense: I call it a *transcendental* illusion because it concerns *the conditions of the possibility of our cognition*. But nevertheless it is only an illusion. Like the vanishing point is only a *focus imaginarius*.

*Ego cogito? I think? No, I only think that I think. It thinks. Thinking thinks. You do not believe it? Think about it! We act, as if there would be an actor, called "I". But in fact, the I is thinkers' fiction. Take a conspicuous example, from what happens in cognition: every morning you see the sun rising, but it is only a practical illusion. We learned from Copernicus: not the sun, but the earth moves. You see: as the sun "rises" in the morning and ends the night, so does our Ego "wake up" in the morning and end our dreams. But as the sun nevertheless remains in the centre of our solar system up there in the silent night of the universe,*

Ego's waking itself is only a dream, a butterfly-dream of the unity of human personhood.

#### I.4 The Christian God and Pope Benedict: Monotheism—Monism—Universalism

A little excursion will point out the dangers of universalism. Our subject today is: Kant and the unity of personhood. I tried to sketch Kantianism as a secularized form of Christianity. I tried to uncover some connecting passages between Kant's concepts of Ego and person and its supposed "unity" on the one hand and the Christian God and its "uniqueness" on the other hand. I underlined the necessity of a metacritique of Kant's criticism and especially of his critique of practical reason. But I didn't give an answer to the question: "Why is such a critique *in concreto* necessary?"

I already told you: I believe such a metacritique is necessary, because a dogmatic uncritical critique of practical reason à la Kant can be dangerous. But why? More concretely speaking: I believe that the faith in the universalization of my will, claimed in Kant's categorical imperative and based on the unity of my person, can be dangerous. We'll see.

I'd like to take a run-up again and start with the crux of universalism in Christianity in order to jump to the crucial problem of universalization in Kant's categorical imperative. Western culture is deeply rooted in Christianity. Christianity is, like the Islam and the Jewish religion, a monotheistic religion. Monotheism is a form of monism (from Greek: *monos*, one). Monotheism is the faith in one God as the only true one. Monotheism is dangerous; dangerous insofar as it claims to have a monopoly on truth. This involves the danger of dogmatism and fundamentalism.

But in reality there is not only one true religion, *una vera religio*, as Augustine defined Christianity. There is not only one God, not only one "chosen people" and not only one "God's country", as many Americans believe. Nobody has a monopoly on truth. There are many equally true, unique religions. Untrue are only those ones that pretend to be the only true ones.

If two such monotheistic religions with the claim to the sole representation of truth are confronted with each other, then there is the danger of a clash. Think, for example, of the continuous war between Israel and Palestine. Monotheism is dangerous insofar as it has a missionary

impact and may easily lead to fundamentalism. When I think of Christianity, I think first of the trails of blood its crusades left behind over the last 2000 years. I think of the crusades at the beginning of the last millennium (1096–1291), blessed by pope Urban, when about 5 million people lost their lives.

Monism and universalism are two sides of the same coin (The Latin *unus* is the translation of the Greek *monos*, one). The Christian claim for universality is based on the “universality of truth”. Pope Benedict confirms “the obligation to send all peoples in the whole world into the school of Jesus Christ, because he is truth personified and therefore the way of personhood.”<sup>2</sup> Dixit Benedictus, the *pontifex maximus* of the *una vera sancta ecclesia*.

Benedict believes in Jesus Christ as “the only saviour of all men.”<sup>3</sup> I believe the world would be saved if it would be saved from saviors who presume to be the only saviors of all men. Jesus Christ may be the savior of many people; this is a faith one may have or not have. By the way: I don’t have it. Thank God I’m an atheist. For faith in Jesus Christ as the only savior of all men is a dangerous superstition.

The light of Christianity is waning, but as the sun going down at sunset it throws long and dark shadows. The evil missionary enthusiasm and militant moralism among Christian fundamentalists like the American so called “new born Christians”, the “evangelicals”, is still alive and was strong enough for a new bloody crusade at the beginning of our millennium. I think of Number 43 and the American crusade against the “axis of evil”—for God’s sake. I talk about the unlawful war against Iraq in which about one million civilians were killed during the last six years. “Collateral damage”? “Mission accomplished”? This is what I called a militant Christian universalism. The creed of love, but indeed: hate and violence. “An ihren Früchten sollt ihr sie erkennen.” So far so bad.

But what about Kant and the dangers of his moral universalism? Modern enlightened people do not believe any longer in the will of an almighty good God (*deus benignus*); in the face of reality one rather could believe in the evil will of a *deus malignus*. Modern enlightened people do not pray any longer: *fiat voluntas tua*. They rather believe in their own free will and as Kantians in their own good will. *Ego cogito*

---

2 Joseph Ratzinger, *Glaube—Wahrheit—Toleranz, Das Christentum und die Weltreligionen* (Freiburg: Herder, 2004), 3. Aufl., 55 f.

3 Ratzinger, 44 and 90.

*et volo*. The *una vera systema* of Kant's *philosophia practica universalis* is a secularized form of the *una vera religio* and its universal pretence.

God's universality and omnipotence is to be found in a secularized disguise in modern moral philosophy, first of all in Kant's categorical imperative to universalize my own will. Was Schopenhauer not right when he said the categorical imperative finally results from a calculated egoism? Let us go a bit deeper into it.

Before I do that, allow me a marginal note on *the* representative of modern Kantianism in Germany: Jürgen Habermas, the most famous German philosopher at the present time, our policeman of moral correctness. In Habermas' "universal morality of enlightenment" we recognize the same problem of universalism. Habermas' universal pretence is the secularized heritage of the Christian claim for universalism; Habermas adopted it from Kant and "de-transcendentalized" it, to use his own word.<sup>4</sup>

Despite the gap between the pre-modern pope Benedict and the modern philosopher Habermas: the claim for universality shows their spiritual affinity. I hold the view that universalism is dangerous, whichever. We should take good care, that we do not universalize our own morality, as the only true one and as the superior one. I quote a sentence I already quoted several times at different places, because it is very telling. Habermas said: "that our Western European morality of abstract justice is developmentally superior [sic! G. W.] to the ethics of any culture lacking universal principles."<sup>5</sup> You see: the best comes from the West.

### I.5 Kant's Ego and its Good—that is, Universally Valid—Will

What is morality? You know: according to Kant *the* principle of morality is the categorical imperative. It is the *one* basic moral law of what he calls pure practical reason. According to Kant, the categorical imperative

---

4 Jürgen Habermas, *Kommunikatives Handeln und detranszendentalisierte Vernunft* (Stuttgart: Reclam, 2001).

5 Jürgen Habermas, quoted in H. L. Dreyfus and S. E. Dreyfus, "What is Morality? A Phenomenological Account of the Development of Ethical Experience", in D. M. Rasmussen (ed.), *Universalism vs. Communitarianism* (London: MIT Press, 1990), 251.

is the only true expression of freedom and freedom means for him the autonomy of will.

The categorical imperative goes (5:30): “Act so, that the maxim of your will may at the same time be regarded as the principle of a universal legislation.” To simplify Kant’s iron rule a bit: beware of the potential *universality* of the principle of your actions. Only if you do that can your will be a good will, because “good” means for Kant: “allgemeingültig”, universally valid. Please mind: Kant speaks *expressis verbis* of *universalitas*, not only of *generalitas*. If you act with a good (i. e., universally applicable) will, you act dutifully and your action can be called moral. That is, in short, Kant’s so-called universal morality of pure practical reason.

Another instructive version of the categorical imperative reads (4:421): “Act so, that the maxim of your action should—by your will—become a *universal law of nature*.” Here Kant’s idea becomes clearer: the parallel of the *law of morality* with the *law of nature*. Even moral philosophy must go the “secure course of science” (Bviif). This was Kant’s dream in the dogmatic slumber of his rational moral philosophy, *more geometrico*. Like every law of nature Kant’s law of morality is characterized by universality and non-self-contradiction.

Let’s recapitulate to be clear and distinct. According to Kant the moral question par excellence is: can you universalize the maxim of your will to act? This ability to universalize is what Kant calls the *positive* criterion of moral actions. Strictly speaking there is a second criterion; Kant calls it the *negative* one. This negative criterion is self-contradiction. Both together are *necessary* and *sufficient* if your will with its maxim is to pass the test of morality. That means, the test-question is: can you universalize the maxim of your will without self-contradiction?

Sounds a bit complicated. An example may clarify. Take a command that is obeyed every day, everywhere in our warlike world, the command to kill. If we would universalize the maxim to kill, this would finally lead to mass extermination and the final consequence would be my own death. With regard to human life the result would be: no more human beings on earth. Therefore this maxim does not pass the test of morality, for Kant. But does it really not pass this test? Let us have a closer look.

The total destruction of mankind may be deplorable for a philanthropist like Kant, to whom mankind was sacred. But—and please pay attention to the following argument—the universalization of the maxim to kill does *not* at all lead to logical self-contradiction! Already Hegel criticized Kant’s “empty moral formalism” and asked: “Where

is the logical contradiction in the idea that there are no more humans on earth?”<sup>6</sup> I repeat: the total destruction of human life may contradict Kant’s idea of the holiness of mankind. It leads to *real self-destruction*, but it does not lead to *logical self-contradiction*.

Conclusion: the maxim “you shall kill” can be regarded as a principle of universal legislation without leading to self-contradiction, and that means this—admittedly misanthropist—maxim is indeed in accordance with the categorical imperative! Is this not remarkable: the total extermination of mankind in accordance with the basic moral law of Kant’s so called pure practical reason!? No misunderstanding: if we only rely on *logic* in a moral philosophy *more scientifico*, as Kant originally intended, then the categorical imperative is not tenable as the basic moral law.

The categorical imperative can only be saved if we take refuge in *religion*. This is exactly what Kant does when he supposes the “holiness of mankind”. He says (5:86): “The moral law is holy ... Man is unholy enough, but the mankind in his person must be holy for him.” Really? At best for heaven’s sake! But for the sake of our earth, it could turn out one dooms-day that it would be best if mankind vanishes, vanishes like a face in the sand next to the sea, “comme à la limite de la mer un visage de sable”, as Foucault said in the last words of his book *Les Mots et les Choses*.

Nevertheless, Kant still believed in the “holiness of mankind”. He emphasized (5:129): “thus the moral law ... leads to religion, that is to *the realization of all duties as commandments of God*.” God comes back through the back door. God becomes a *postulate* of practical reason. *Deus ex moralitate*. God as a “*postulant*” in the church of Kant’s moral universe. In transcendental moral philosophy no answer is possible to the question, *if* God really exists, and with regard to Kantian moral philosophy such an answer is not necessary. It is only necessary that we—as good Kantians—act, *as if* god exists. But only pretending God’s existence: is that not a form of transcendental hypocrisy?

---

6 Hegel, *Grundlinien der Philosophie des Rechts*, §135.



## I.6 Moral Progress and the French Revolution

Well, let us forget about the “transcendental ideal”; let us forget about God and his dubious existence. Let us come down to earth again. Welcome to reality and the alleged “holy mankind”. Kant believed the destiny of mankind is an incessant *progress*. According to Kant, pure practical reason entitles us to presume that the world on the whole always progresses for the better. I doubt it! Hope for progress, especially hope for moral progress, is dope, Pandora’s dope. It produces sweet dreams instead of facing reality. I agree with John Gray, who recently showed in his book, *Straw Dogs*, that the hope of progress is an illusion: “As the hope for a better world has grown, so has mass murder.”<sup>7</sup>

One day, when the “totally enlightened earth will shine in the sign of triumphant disaster”, as Adorno said shortly before the end of the second world war in his *Dialektik der Aufklärung*,<sup>8</sup> humans will be burned like Laozi’s “straw dogs”<sup>9</sup>—like the victims in Hiroshima and Nagasaki. Anyway, let us not be unfair: Kant had no chance to compare his sweet dream of eternal peace with the nightmare of two world wars and weapons of mass destruction in the twentieth century. But in his late essay *Der Streit der Fakultäten*, from the end of the eighteenth century (1798), Kant refers to *the* outstanding historical event of his time, the French Revolution: in his opinion, it proved the moral tendency of mankind. (More precisely, he speaks of the way of thinking of the spectators of the French Revolution. I am sure he did not mean the spectators of the executions during the years 1793–1794.)

The historical background of Kant’s “revolution of thinking”, especially in his moral philosophy, is the French Revolution. As for the *terreur* of the French Revolution, I remind you of the chapter, “Absolute Freedom and Terror”, in Hegel’s *Phänomenologie des Geistes*, where he demonstrated the dialectics of moral fundamentalism. Listen to the following words and guess whose words they are: “We want to substitute morality for egoism, principles for habits, duty for propriety and the power of reason for the obligation of tradition.”<sup>10</sup> These are not the

7 John Gray, *Straw Dogs* (London: Granta Books, 2002), 96.

8 Theodor W. Adorno, *Dialektik der Aufklärung*, (Amsterdam: Querido, 1955), 13.

9 Laozi, ch. 5.

10 M. de Robespierre, quoted in H. Mainusch, “Auf dem Weg zu einem Weltethos”, in *Dao in China und im Westen* (Bonn: Bouvier Verlag, 1999), 200. My translation.

words of Kant as I supposed, when I read them first, but the words of his contemporary Maximilien de Robespierre, the executioner of the French Revolution, a man who said he would even use terror as a means to achieve virtue. And he really did this, until his reign of “morality” (that turned out to be a reign of terror) came to an end in 1794, and he himself was executed on the guillotine.

Moralistic purism, related to fundamentalism, has a predisposition to turn into moral terrorism. They are two sides of one coin. And the categorical imperative of pure rational morality has a predisposition to become its own opposite. This is its dialectic.

### I.7 Duty Ethics and its Dialectics

You remember: To act according to the categorical imperative means to act with a good, that is a universally applicable, will. If we do this, our acts are purely rational and not emotional. In other Kantian words: we act only *dutifully* (*aus Pflicht*), not out of inclination or affection (*aus Neigung*). I always found Kant’s histrionic exclamation suspicious (5:86): “Duty! You great lofty name ... you demand submission ...” Duties as commandments of God excite a holy shudder in Kant’s soul—*wie schauerlich!*

Following Kant, respect for authorities (“Gehorchet der Obrigkeit”) is a categorical imperative. Because “all authorities are from God”—oh my God!—for a subject no resistance is allowed, not even against a tyrant. (I’ll come back to this dark point when I say something about the Mengzi.) Following Kant, the only thing the subject has to do is to obey. The trial to kill the monarch because of tyranny is high treason and must be punished with the death penalty, as Kant says. (*Nota bene*: Kant was an advocate of the death penalty. The only just answer to murder for him was the death-penalty. He called it the categorical imperative of jurisdiction and condemned the *compassibilitas* of opponents as an “affected humanity”.) The resistance against Hitler would not have found Kant’s approval.

I repeat: unquestioning respect for authorities is a categorical imperative for Kant, because resistance against the highest power would be a self-contradiction. (If such resistance would be successful, there would be a higher power than the highest power.) Here you see how the categorical imperative can be used, or misused, to protect tyranny. By the way: Arthur Seyss-Inquart, who was executed in 1946 as one of the ten

main guilty persons of the NS-regime, said at the court in Nürnberg, where the trials against the Nazis took place: “The number of human beings you can kill on the basis of hate or fancy for massacre is limited, but the number of those you can kill cold-bloodedly and systematically in the name of a military ‘categorical imperative’ is unlimited.”<sup>11</sup> Is this a malicious distortion of Kant’s categorical imperative? No, I don’t think so, although what he is talking about is not a categorical imperative *sensu strictu*.

I have to add here some very serious and alarming sentences from chapter VIII, “Duties of a Law-Abiding Citizen”, in Hannah Arendt’s book, *Eichmann in Jerusalem, A Report on the Banality of Evil*:

... during the police examination ... he [Eichmann] suddenly declared with great emphasis that he had lived his whole life according to Kant’s moral precepts, and especially according to a Kantian definition of duty ... And to the surprise of everybody, Eichmann came up with an approximately correct definition of the categorical imperative ... Upon further questioning, he added that he had read Kant’s *Critique of Practical Reason*.<sup>12</sup>

After some critical remarks on Eichmann’s (mis)understanding of Kant, Arendt concludes:

Whatever Kant’s role in the formation of “the little man’s” mentality in Germany may have been, there is not the slightest doubt that in one respect Eichmann did indeed follow Kant’s precepts: a law was a law, there could be no exceptions ... No exceptions—this was the proof that he had always acted against his “inclinations” [*Neigungen*], whether they were sentimental or inspired by interest, that he had always done his “duty”.<sup>13</sup>

Is this a diabolical distortion of Kant’s duty-ethics? I think it is the dialectics of duty-ethics. The “radical good” has the intrinsic tendency to become its very opposite, the “radical evil” or the “banal evil”.

---

11 Werner Stegmaier, zitiert nach Mainusch, “Auf dem Weg zu einem Welteithos”, 205. Übersetzung aus dem Deutschen vom Verfasser.

12 Hannah Arendt, *Eichmann in Jerusalem* (New York: Penguin Books, 1963), 135 f.

13 Arendt, 137. Compare the German version: Hannah Arendt, *Eichmann in Jerusalem, Ein Bericht über die Banalität des Bösen*, 3. Aufl. (München/Zürich: Piper, 2008), 231 f. I owe these references to Stephen Palmquist and Wong Kwok Kui.

Maybe Nietzsche had a good nose when he said: “The categorical imperative smacks of cruelty.”<sup>14</sup> My metacritical conclusion: I believe it is our unconditional duty to be very skeptical in view of unconditional duties. My “categorical imperative”: no categorical imperatives!

### I.8 *Ratio et emotio*—Kant on Compassion

Let’s recapitulate one last time before we go East. Kant’s pure practical reason commands us to act *only dutifully*, not out of inclination or affection. In other words: pure practical reason commands us to act *only rationally*, not emotionally, because only then is the potential universality of the principle of our action guaranteed.

I vary a famous word of Kant and admit: emotion without rationality is *blind*. But I object: rationality without emotion is *empty*. *Homo sapiens est animal rationale et emotionale*. A radical “emotivism” is as blind as a radical moral rationalism à la Kant is empty. It is an “empty formalism” without any protection against abuse.

I have tried to show that pure practical reason is dangerous, because obeying the categorical imperative as one’s basic moral law may have disastrous consequences. As I said at the beginning of my paper: Kant’s categorical imperative, his so-called practical reason and his good will, is not a good guide on our way to practice. It is unsuitable to regulate moral action. It is practical only in theory, not in practice. If we try to go along with it in practice, it may easily lead to our fall.

As Wittgenstein the second, the Post-Kantian, said in his *Philosophical Investigations*: “We have got on to slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk, so we need friction. Back to the rough ground!”<sup>15</sup> Friction in a moral sense is affection, emotion, sentiment. Hume was right: morality is determined by sentiment. But Kant kept on dreaming the sweet dream of good will in his dogmatic slumber of pure practical reason.

---

14 Friedrich Nietzsche, *Zur Genealogie der Moral*, Part 2, §6, in his *Sämtliche Werke, Kritische Studienausgabe*, ed. Giorgio Colli and Mazzino Montinari (Berlin: Walter de Gruyter, 1980), vol. 5, 300.

15 Ludwig Wittgenstein, *Philosophical Investigations*, tr. G. E. M. Anscombe, second edition (Oxford: Blackwell, 1958), 46 (§107).

Before Kant it was Hume, and after Kant it was Schopenhauer, who knew that ethics does not work without sympathy and compassion. Schopenhauer knew that abstract concepts *a priori* never can motivate humans to act morally. But for Kant, devoted only to cold abstract duty, the warm feeling of compassion is weak and always blind. He admitted in *Anthropologie in pragmatischer Hinsicht* that a suffering child may fill our heart with melancholy and compassion, but he adds that this compassion never has the dignity of virtue. For Kant compassion is an affect and affects are pathological attacks that interfere with our freedom and impair the self-control. For Kant compassion is effeminate and childish. No comment—only compassion.

But, dear colleagues, please imagine a playing child who is about to fall into a well. Next to the well sits a Kantian, re-reading his *Critique of Practical Reason*. He just reads and re-reads the passage: “Act so, that the maxim of your will ...”. He sits and reflects ... and then we hear a splash. I’ll come back to the poor Chinese child in a minute. Meanwhile, mind: it is not reasonable to be only reasonable.

I’ll go back now, from Kant’s iron rule to Mengzi, via Kongzi’s golden rule. I’ll go back with a remarkable passage in Kant’s little book *Grundlegung zur Metaphysik der Sitten*, where he quotes the golden rule; not Kongzi’s version, but the nearly identical Latin version. He calls it trivial and unsuitable as a guiding moral principle. Why? It cannot be a universal law, because it does not contain the reason for the *duties* human beings have. Now I feel a strong affection for leaving Kant and going back to Asia.

## II. The East: Kongzi, Mengzi, Huainanzi, Zhuangzi

### II.1 Kongzi’s Golden Rule

The Roman Emperor Alexander Severus was so delighted with the *regula aurea* (i. e., the golden rule), that he posted it on his palace: *Quod tibi fieri non vis, alteri ne feceris*. Translated into a well known German rhyme: “Was du nicht willst, das man dir tu, das füg auch keinem andern zu.” “What you do not wish to be done to yourself, do not do to others.”

The oldest secure reference to the golden rule in the West is to be found in ancient Greece in the words of Isokrates (436–338 BC). We also find it in the Old Testament (Tobias 4:15a) and in the Sermon on

the Mount in the New Testament (Matthew 7:12 and Luke 6.31). Different versions are to be found in Islam, Buddhism and Hinduism.

If there would be a universal moral rule, then it would be the golden rule. The *locus classicus* of the *regula aurea*, the most prominent and probably the oldest reference of this world-wide ethical law, is in Kongzi's *Analects*: "Is there any single saying that one can act upon all day and everyday? The Master said: Perhaps the saying about consideration: 'never do to others what you would not like them to do to you.'"<sup>16</sup>

The one word that will keep us on the path to the end of our days is *shu*. Waley translates it as "consideration". Others use "reciprocal consideration". I prefer "reciprocal empathy". *Mathews' Chinese-English Dictionary* (no.5875) reads: "*shu*, the principle of reciprocity, making our own feelings [sic! G. W.] a rule whereby we are guided in dealing with others." Indeed, the crucial point seems to be that the word *shu*—with the radical *xin*—concerns primarily our own feelings and empathy with others.<sup>17</sup>

Originally it is not (only) aimed at our (*head-*)mind, at our reason and self-reflection. It is rather aimed at our *heart-*mind (*xin*) and compassion, our empathy and sympathy. Kongzi's golden rule basically is not an abstract rational operation of universalizing my own will or the intentions of my own Ego. It is not something like a half-baked categorical imperative as some German philosophers and sinologists (e. g., G. Paul and H. Roetz) have it. Kongzi's golden rule deals with reciprocal empathy and *responsibility* in the literal sense of this word, having to do with *response* and being responsive to somebody in a concrete situation. The Chinese character *ren*, humanity, shows that humanity has to do with *two* people.

I understand the golden rule as a verbal formulation of a preverbal, intuitive way of sympathetic resonance, as the verbalization of an archaic reciprocal resonance. I'll come back to it shortly. But what we already see: what a difference to Kant! The Neo-Kantian attempt to kantianize Kongzi is ideological Neo-Colonialism or philosophical figurism. I recommend to those Neo-Kantians: back to Rousseau! He emphasized, correctly, that the real fundamentals of the golden rule are conscience and feeling. This applies exactly to Kongzi's words. In my opinion,

---

16 Kongzi, *Lunyu*, tr. Arthur Waley (Ware, Hertfordshire: Wordsworth Editions Ltd., 1996), 15.23 (207).

17 Cf. U. Unger, *Goldene Regel und Konfuzianismus*, in *Minima Sinica*, ed. Wolfgang Kubin and Suizi Zhang-Kubin (Bonn: Edition Global, 2003), 2, 19–41.

the true path does not lead from Kongzi to Kant, but the other way round! Via Rousseau it leads from Kant back to good old Kongzi.

## II.2 Mengzi on Regicide and on Compassion

There are different versions of the golden rule in the *Mengzi*. I quote a passage from *Mengzi* 4 A9, showing the political relevance of the golden rule. Legge translates: “There is a way to get the people: get their hearts, and the people are got. There is a way to get their hearts: it is simple to collect for them what they like, and not to lay on them what they dislike.”<sup>18</sup> The point in the *Mengzi* is: the golden rule is true of ordinary people as well as of the reigning monarch. What he forbids the people to do, he himself also should not do. I remind you of the famous passage in *Mengzi* 1B8:

“Is regicide right?”

“He who outrages benevolence is an outrager,

He who outrages righteousness is a cruel fellow.

He who is a man both cruel and outrageous is a despot forsaken by all. I have heard that Zhou, the despot forsaken by all was killed, but not that the killing was regicide.”<sup>19</sup>

The political consequences of Mengzi’s understanding of *shu*, of reciprocal empathy, and the political consequences of Kant’s understanding of his categorical imperative show clearly that their ways of thinking lead into different directions. But the differences between Kant and Mengzi do not only concern despots. It also concerns everyday life; it concerns, for example, compassion in dealing with other persons, with children, and even with animals. For Kant’s rational intelligence of *head-mind*, compassion was an “affect”, “weak and blind”, lacking the “dignity of virtue”. For Mengzi, compassion is the *punctum saliens*, the most important point of humanity. For him it is the heart-mind (*xin*) that thinks (*si*).<sup>20</sup> (*Nota bene*: the radical *xin* is also a part of the character *si*.)

---

18 James Legge, *The Chinese Classics*, vol.II, *The Works of Mencius* (Taipei: SMC Publishing Inc., 1991), 300.

19 Mencius, *Library of Chinese Classics* (Human: People’s Publishing House, 1999), 43.

20 Mengzi, 6 A15.

Mengzi's emotional intelligence reminds more of Pascal's *raison du coeur* than of Kant's "reine Vernunft". For Mengzi, all people have a heart-mind that "cannot bear" (*bu ren*) to see the sufferings of others.<sup>21</sup> "When I say that all men have a (heart-)mind which *cannot bear* to see the sufferings of others, my meaning may be illustrated thus: Even nowadays, if men suddenly see a child about to fall into a well, they will without exception experience a feeling of alarm and distress."<sup>22</sup> Usually the *bu ren*, "cannot bear", is translated by "compassion". The word *ren* is instructive: the character has a knife over the heart. Please mind: Kongzi's *shu*, "reciprocal empathy", as well as Mengzi's (*bu ren*), "compassion", have the radical *xin*, "heart-mind". Isn't that telling? Last but not least, the double meaning of the character *xin* is itself telling: heart and mind—for Kant a contradiction.

In contrast to Kant, for whom animals are merely things, objects one can treat as one likes, I finally quote a last passage from the *Mengzi* 1 A7, where *bu ren* occurs again.

The king ... was sitting aloft in the hall, when a man appeared, leading an ox past the lower part of it. The king saw him and asked: "Where is the ox going?"

The man replied: "We are going to consecrate a bell with its blood." The king said: "Let it go. I cannot bear [*bu ren*] its frightened appearance, as if it were an innocent person going to the place of death."<sup>23</sup>

Kant would have condemned this *compassibilitas* as sentimentality. Well, *tempus fugit*. Last but not least, let us have a short look at the *Huainanzi* and the *Zhuangzi* and let me concentrate on the character *ying*, for it also has the radical *xin*.

### II.3 Huainanzi's Reciprocal Resonance: *Ying* and *Ganying*

The concepts of *ying* and *ganying* are basic concepts in East-Asian ethics. Especially in Daoist "ethics", an ethos without morality, these concepts play an outstanding role. What does *ying* mean? In *Mathews' Chinese-English Dictionary* (no.7477) we read that *ying* means: "ought, should, must; suitable, right, fitting; necessary etc." With a different intonation it means: "to reply, to respond, to echo, to correspond etc." And *ganying*

21 Mengzi, 2 A6.

22 Legge, 202.

23 Legge, 139.



*ing* (no.3232) means: “moved to response through the feelings and affections: induction”.

Charles le Blanc, the translator of the *Huainanzi*, emphasized that it was around the idea of resonance (*ganying*) that the philosophy of the *Huainanzi* was elaborated. Chapter VI is devoted explicitly to the idea of resonance.<sup>24</sup> The origin of the notion of *ying* is an acoustic and musical one. Chapter VI, 6b of the *Huainanzi* gives a famous example of spontaneous reciprocal resonance, the example of the two zithers: “When the lute-tuner strikes the *kung* (*gong*) note (on one instrument), the *kung* note (on the other instrument) responds (*ying*) ... This results from having corresponding musical notes in mutual harmony.” The two instruments cor-responding, responding to each other reciprocally (*xiang ying*) and being in mutual harmony, are like Zhuangzi’s fishes enjoying each other and the heart-minds of people who vibrate in reciprocal resonance and sympathy.

By the way, the Greek word *sympathy* originally meant something like a sympathetic vibration, for example of bronze vessels. In music it was used of chords that vibrate together. The heart-minds of these people who feel sympathy react in reciprocal spontaneity (*xiang ran*) without acting on purpose. They even forget each other (*xiang wang*) and respond in reciprocal oblivion, in reciprocal resonance and concordance. This intuitive spontaneous reciprocity and empathy results by itself (*ziran*) without any selfishness, without any rational reflection and without moral principles. This inductive, sym-pathetic, com-pasionate cor-respondance results from a natural feeling of “responsibility” that is beyond morality and before morality in the sense of moral principles, laws, and duties. It is a *resonance du coeur*, a resonance of the heart-mind without reasoning.

Is all this too harmonious, too idyllic, too romantic? I don’t think so. This spontaneous natural resonance follows the spontaneous necessity to do what has to be done. Please remember that *ying* also means “necessary, fitting, and suitable. *Ying* means the necessary fitting response according to the changing situation. Right and wrong are situational. In the appropriate situation nothing is wrong. Without the ap-

---

24 Charles le Blanc, *Huai-Nan-Tzu, Philosophical Synthesis in Early Han Thought* (Hong Kong: Hong Kong University Press, 1985), 9.

propriate situation, nothing is right.”<sup>25</sup> In the *Liezi* we read the remarkable words:

In any case, nowhere is there a principle which is right in all circumstances, or an action that is wrong in all circumstances. The method we used yesterday we may discard today and use again in future, there are no fixed right and wrong to decide whether we use it or not.<sup>26</sup>

The highest principle is without principles. The one who acts according to this principle without principles acts primarily without rational reflection. First of all he only reflects (i. e., mirrors the actual situation). Speculation and re-flexion re-sult out of it.

In chapter VI, 6b of the *Huainanzi*, in a context where the author talks about *xiang ying*, reciprocal resonance, he gives the interesting example of a burning mirror that gathers fire from the sun. Probably this burning-mirror was a concave bronze-mirror that was used in Zhuangzi's days to focus the sunrays in order to set fire.<sup>27</sup> Is the “response” *ying* of this burning mirror not a very practical and striking example for the Daoist *wei wuwei* or *wuwei ziran*? It is an example of doing something without interference of myself, only by responsive self-so-ing (*ziran*). It is an example of “speculation” (from Latin *speculum*, mirror) without intellectual speculation. I come to the end with a little remark on the *Zhuangzi*.

#### II.4 Zhuangzi's Ethos without Ego: The Empty Mirror

Looking for older sources of the Daoist key-term *ying*, we have to go back to the *Zhuangzi*. In the “miscellaneous” chapter, 33.5, we find a saying that I would like to call the *minima moralia daoistica*. It deals with the “true man of the *dao*”. It reads: “His movement is like water, his stillness is like a mirror, his response (*ying*) is like an echo.”<sup>28</sup> The end of the “seven inner chapters” (chapter 7.6) says:

25 *The Tao of Politics—Lessons of the Master of Huainan*, tr. T. Cleary (Kuala Lumpur: Konsep Lagenda Sdn Bhd, 1992), 39.

26 *The Book of Lieh-tzu*, tr. A. C. Graham (New York: Columbia University Press, 1990), 163.

27 Cf. H. H. Oshima, *A Metaphysical Analysis of the Concept of Mind in the Chuang-Tzu*, in V. H. Mair (ed.) *Experimental Essays on Chuang-tzu* (Honolulu: University of Hawaii Press, 1983), 63 f.

28 *Wandering on the Way*, tr. V. H. Mair (New York: Bantam Books, 1994), 341.

“Just be empty, that’s all. The mind of the ultimate man functions like a mirror. It neither sends off nor welcomes, it responds (*ying*) but does not retain .”<sup>29</sup>

But how can the heart-mind of a true human being respond like a mirror? We already heard it: one must be empty; one must be without an Ego. The first chapter (ch. 1.1) of the *Zhuangzi* closes with the path-breaking words: “... the ultimate man has no self (*wu ji*). The spiritual person has no accomplishment, and the sage has no name.”<sup>30</sup> No name, no fame—no ego.

I agree with Feng Youlan, who said that to be without a self is the essence of Zhuangzi’s philosophy.<sup>31</sup> The true human being of the *dao*, who follows nature (*ziran*—self-so), has no personal self. Such a human being is not a *person*. I *personally* believe this is true.

## II.5 Final Remark: Mirror-Neurons

What I want to say can be expressed in one sentence: the Daoist ethos without morality, this ethos of reciprocal resonance and pre-rational reflection, pulls us from our top-heavy head back on to our feet and enables us to walk on the “way”. Please do not think the old Daoist stories of mirror-reflection are outdated. On the contrary!

One of the latest hits in neuro-biology is: *mirror-neurons*. The Indian neuro-physiologist Ramachandran, as far as I know the discoverer of these mirror-neurons, calls them empathy-cells and claims they are the neurological basis of ethics and morality. Mirror-neurons are spontaneously and unintentionally activated in the heart-mind of passive observers as if they had themselves acted. Without mirror-neurons there is no empathy, no sympathy, no compassion. Mirror-neurons are responsible for pre-rational, spontaneous intuition. Mirror-neurons are responsible for our “emotional intelligence”. Some scientists have called this communicative resonance an “intersubjectivity without subjects”.<sup>32</sup>

---

29 *Wandering on the Way*, 71.

30 *Wandering on the Way*, 5 f.

31 Cf. *Chuang-Tzu, A New Selected Translation* by Yu-Lan Fung, (New York: Paragon, 1964), 81.

32 Cf. J. Bauer, *Warum ich fühle was du fühlst* (München: Heyne, 2005), 63.

What a pity that these new neuro-biologists and philosophers do not know that they are neo-daoists.<sup>33</sup>

---

33 During a little radio-interview before the beginning of the conference the interviewer asked me: “Who are you?” I replied. “I’m the German scholar.” “Oh, I understand”, he said, “you are here to make the conference more sexy.” I am still wondering if he was right. The more detailed version of this paper is to be found in my last little book, *Splitter—Spitze Bemerkungen zu Kant und dem Käntchen Jürgen Habermas*, available for free download on my home-page: [www.guenter-wohlfart.de/books](http://www.guenter-wohlfart.de/books).

## *Keynote Essay to Book Three*

### Incorporating Kantian Good Will: On Confucian *Ren* (仁) as Perfect Duty

*Chung-ying Cheng*

#### 1. Introductory Remarks

On the basis of autonomy of good will, Kant is able to transform virtues of good character into duties of right action according to moral rules of action that embody good will and obey the categorical imperative. He sees the necessity for realization of virtue in action to be intrinsically motivated by good will and moral law. Thus the ultimate purpose for deontology is not only to rationalize actions of virtues under autonomy of good will but to assure universality and universalization in practice. With this purpose in mind, four kinds of duties arise for justification: two “perfect” duties, respectively to oneself and to others; and two “imperfect” duties, respectively to oneself and others. Finally, Kant introduces the task of achieving the supreme good as a challenge to the performance of moral duties, and this inevitably leads to his “antinomy of practical reason”. He resolves this antinomy by appealing to the postulates of God and immortality of soul.

I have two purposes for this essay: one general and one specific. My general purpose is to question how cogent and solid is the metaphysical foundation of the Kantian moral system. This will be argued in the sequel, to be published elsewhere. Although Kant’s moral philosophy of good will and moral law appears to be well argued and well formulated, with a comprehensive scope, it nevertheless suffers from three fundamental drawbacks that should force him to look for some basic revision and broad solutions. As an adequate response to these three drawbacks, Confucian moral philosophy based on the experienced nature of humanity should come to the rescue as both a new foundation and an

ideal end that resolves the Kantian antinomy of practical reason in a more coherent manner.

Concerning Kant's distinction between perfect and imperfect duties, questions are raised about their amphibolous validity and meaningful application. For duty is what obligates one to act in accord with the moral law, whether inward or outward. As modalities of virtue-duties, instead of speaking of perfect duties and imperfect duties, these terms being misnomers by amphiboly, we should speak of duties to be performed under different modes (such as necessity, contingency, and possibility) that impact the ways we perform duties as human persons. Hence we can speak of necessary duties, contingent duties, and possible duties according to modes of time and situation. Perhaps Kant wishes to identify necessary performance with "perfect", relativity and contingency with "imperfect". On my suggested modes of duties, however, Kant could allow possible duties to subhuman beings (such as animals) or superhuman beings (such as spirits and God or Heaven) that he would otherwise regard as simply "imperfect". One must remember that possibilities could be realized as temporary contingencies and even as temporarilized necessities.

In my analysis of reasons for the four Kantian duties, to be presented in the sequel, we shall see that two such duties are actually derived from a basic underlying implicit duty of life-respect, and the other two are actually derived from a basic underlying implicit duty of society-respect. This deontological derivation conveys both a sense of presupposition and a sense of disclosure in dutification of the two virtues of life and two virtues of society. By introducing underlying duties of life-respect and society-respect, we may arrive at a substantive principle of law that is essentially the principle of *ren* (仁, benevolence/beneficence), combining the implicit duty of life-respect and the implicit duty of society-respect in a deep experience of the human person. From this Confucian perspective, Kant's "imperfect" duties, like his perfect duties, are perfect parts of the unavoidable duty toward life and society. They are all necessary and hence "perfect" duties under normal circumstances of life and society for their preservation and development. They can become contingent under special conditions when circumspection is required. Based on these considerations, we are able to introduce the Confucian principle of *ren* that incorporates and integrates the Kantian good will into the nature of the human person in an autonomy-creating dynamic unity as given in our deep reflective experience of humanity. This first essay in the series introduces the Confucian *ren* as the perfect duty in the sense of

being necessary and ideally perfect. My goal in this new approach is to anchor, substantiate, and consequently, renovate the Kantian approach; it may be titled a Confucian-Kantian synthesis of the empirical and the *a priori*.

## 2. *Ren* as the Centralizing Moral Principle in *Humanity*

There are four stages of the Confucian development of the concept of *ren* as the central principle of morality in humanity. Let us begin with Confucius's reflection. For Confucius, morality is a matter of humanity as expressed in communal codes of behavior going by the name of *li* (禮, rules of propriety derived from social sentiments for correct action). Two aspects of *li* could be mentioned: the social and communal reference of *li* makes *li* a conventional and yet axiological bonding among people. To offend *li* is to violate the social order and common spirit of community and needs to be stopped. Second, *li* is rooted in the function of communication with the spirit of one's ancestors and the spirits of environments that are established by practice and understanding. With these two aspects it is not necessary that *li* be regarded as a formal and rigid code of behavior. It is to be observed and performed with understanding of others and expectation of understanding of oneself from others. It is an expression of inner feelings and core values across a community with proper form. This means once we find a better form of expression or our common feelings have transcended their given forms, new forms of *li* could evolve and be adopted. In this sense *li* crystallizes a deep wisdom of the human mind and is intended to give individuals and their actions a proper place in the system of ordering and organization of a community. It is therefore also a vehicle for the sustenance of a common sense of rightness in a community. Yet because of this, *li* can be reformed and changed in terms of a sense of right or new demands of fairness and justice; it can be reformed and changed in terms of the deep feelings of humanity under different circumstances and in light of the wisdom and insights of the people who could influence and provide a new order of expression and realization.

What makes *li* possible, therefore, is the realization that deep humanity is to be shared by all persons and that this provides a control and regulation of one's private interests and desires. This realization is the experience and discovery of humanity known as *ren*. It is the deep sense of what is given in a person for this life as the fundamental

reality. Hence Confucius is described in *Analects* 9:1 as one who does not speak of profit, but gives himself to *ren* and *ming* (命, destiny / fate). This is a significant statement, for it tells us that *ren* is an attitude and practice of a conviction based on a vision or understanding. It is not accidental that *ren* is mentioned together with *ming*, as the sense of destiny one has reached in one's life. Since Confucius said he comes to know the *ming* of heaven (*tianming*, the mandate of heaven) at fifty, we may infer it is at about that stage of his life that he achieved deep insight into the meaning and significance of *ren* and devoted his life to practice *ren*. He speaks of *ren* more than 100 times in the *Analects*, identifying and explaining the nature of *ren* and how it can be practiced. This no doubt demonstrates that *ren* must be considered a central principle of his life and faith.

But what is *ren* after all? How is it related to us as human beings? Many explanations of *ren* by Confucius command our attention. Two are basic explanations: the intuitive one is that *ren* is to love people; the analytical one is that *ren* is to discipline oneself and practice *li*.<sup>1</sup> To combine both, we can see how one can come to have *ren* as love of people as a result of controlling one's selfish interests and private desires and practicing the *li*. In this fashion we can see how *ren* comes from inside oneself, not from outside oneself. In other words, it is not just a matter of following the rules of *li* in society but a natural desire and capacity to follow the *li* from one's heart and mind. It has its autonomy not by self-legislation without content, but by self-reflection on experiences so that one's mind and will grow stronger and independent in making one's free choice of what is desirable. There is a strong sense of the power of will as Confucius says in *Analects* 7:29: "If I desire to have *ren*, this *ren* comes to me in no time".

One may ask: how does one decide on the practice of *ren* or follow the relevant rule of action motivated by *ren*? The answer is the famous golden rule in *Analects* 12:2: "Do not do to others what I do not desire others do to me" (己所不欲, 勿施与人). With this rule of action one comes to have a criterion on how to love people and how to *keji* (克己, discipline oneself) and *fuli* (复礼, practice *li*). This rule of course can be said to be equivalent to the demands of the self-legislated law of morality in Kant. But there is a big difference in the formulation of these laws of morality. For Kant, the stress is put on the universality or universalization of the application of a given rule of action, whereas for Confucius

---

1 *Analects*, 12:22 and 12:1.



the stress is put on the desirability and suitability of the content of an action. One requires a form to be observed, the other requires a feeling as the content of moral action to be a guide and judge of what not to do. What one should not do, if the content requirement is not satisfied, is implicitly universally required. The *ren* in the Confucian rule of reciprocity is for individuals in general, even for groups of people. It is even the basis for correct action or non-action toward people in government. This makes this attitude of *ren* closer to the third formulation of Kant's categorical imperative (CI). There is not only autonomy of will here, but also a will that is concerned and cares. It is not just a free will that makes decisions independently, but a will that considers consolidations without losing independence. In this regard the will of *ren* is not only free and open but creative. Through its creativity it is able to face reality as a resource and to integrate experience as a basis or ground for decision-making. But if we bring out the *li* element of what is required, then the *li* functions as maxim and should conform to the requirement of universality of individuals and society together.

This is one important extension of the negative formation of *ren* as holding oneself and performing or practicing *li*. However, there is a positive formulation of *ren* in *Analects* 5:30: if one reaches an end, one should help others to reach their goals. Here we can see that rules can be formed so that we can follow them in helping others, even though one can go ahead to give help to others without any fixed rule but with proper knowledge that would help and identify who needs help most and in what way. What we need to pay attention to is that one can help others both negatively and positively and with knowledge and feelings at the same time. This differs greatly from Kant in the following way: (1) it is more concentrated in experience than Kant; (2) it covers content both positively and negatively; and (3) it can be supported by feelings and desires. Let me explain.

While Kant proposes a formal criterion and requirement regarding the practicality of a maxim, he has no concern about which or what kinds of maxims would fit with the formal criterion. This makes the criterion an empty formula for discovering suitable maxims. If we take the requirement of universality on the basis of experience, it is difficult to see any maxim that would hold universally as we could not go over all cases of its application. The most we can make out is some inductive generalization based on a large number of cases. We have to appeal to our best imagination and intuition to decide which would be genuine maxims of morally good action. But without experience our imagina-

tion and intuition could be blind too. This means the formal criterion presented in the CI could give little help or hope in locating maxims as practical guides. Now with the Confucian concentration on a person's self-feelings and self-desires, each person could discover his or her own maxim of moral action and apply it universally and reciprocally with others. Why? Because one could use oneself as a measure for deciding practically what not to do and what to do. Even if I could be wrong about myself, I would not do harm in terms of what I do not like others do to myself. As to what I could do to others I have to see how others like to be helped themselves just like I would like to see how I myself am to be helped. In this sense I am not to impose on myself and others what to do and what not to do. I have to discover what others like and do not like by discovering what I like and do not like. I have to reflect what I want to do and not to do. Similarly I have to observe what others do not like to do and what they like to do. The negative and positive golden rules rooted in self-feelings and observations of other-feelings make it possible to identify equivalent or same needs and ends in others so that we can act rightly. This amounts to making both the negative and positive golden rules of morality a discovery procedure for potentially each and every action to be done.

We then come to the question whether one can always appeal to one's own feelings in order to find out what others would like to do or not to do to. Theoretically, nothing prevents one from reading one's own mind, but one needs to form the habit actually to do so. It is obvious from experience that many moral decisions are misled or wrong simply because one is not good at reflecting on what one really needs and what one really dislikes. For this reason the *Analects* advise us to practice reflection often. Consider Zengzi's statement in *Analects* 1:4: "I have reflected on myself three times a day: do I do my best in my dealings with others? Do I remain truthful and honest in transactions with friends? Do I review what I have learned from my teachers?" Reflection on what one has learned is revealing: it is not just what one does with regard to others, but one has to deal with oneself in a morally right way, making sure one improves oneself and makes progress in one's abilities on a daily basis. In general a second order moral law based on understanding the importance of self-reflection and observations of others is to make self-reflection and learning the principle of moral efficiency. This can be called the Principle of Self-Cultivation (*xiuji*, 修己; or *xiushen*, 修身). It is a principle that enables us regularly and even systematically to discover first-order rules of moral action. This is because it

requires us constantly to practice reflecting, observing, and learning so we may come to know what are the correct ways of doing and not doing, according to the two golden rules of action: the negative and positive. In this fashion, one can be sure one is always learning from oneself and others—a creative resource given to us by nature.

The Principle of Self-Cultivation is first formulated by Confucius in his statement in *Analects* 14:42, regarding how to become a self-ruling person (*junzi*, 君子): “To cultivate oneself toward respect, cultivate oneself toward settling others, cultivate oneself toward settling the people.” This principle, though general, is not empty because there are ends to be attained even if no specific event has taken place. These general ends are self-respect (*jingji*, 敬己), others-settling (*anren*, 安人) and people-settling (*anbaixing*, 安百姓). So-called “settling” is a matter of bringing harmony and peace to others and people so that there are harmonious and peaceful relationship between me and others and between me and the people, if I am the political leader. This is to cultivate moral order without focusing on any specific action. But it is also a general requirement that makes the application of the golden rules always ready to be used in order to be specifically relevant to concrete occasions of life. This Principle of Self-Cultivation has been explicitly stated by the classic text, the *Great Learning*, and is implicitly assumed by all major Confucian texts from Zisi to Mencius to Xunzi.

We may now query whether this principle of cultivation is a moral law in the Kantian sense. Could it be required by the good will and even self-legislated? My answer is absolutely positive. There is no reason why the good will could not see intrinsic good in requiring itself to cultivate itself as a sovereign by itself and also as a sovereign that rules over feelings and desires of the human person it reigns. It is not only intrinsically good to do so but it will also bring good to others if it is at the same time the pure reason of understanding in its practical use with regard to people and the world situation. In others it could will good that will hold universally. In this we may now formulate this Confucian principle of morality as:

S: Make decisions and act in such a way that you will bring respect to yourself, harmony to others, and peace to people.

On the basis of this principle, one may move to the two golden rules of moral action:

N: Make decisions and act regarding things not to do to others by reflecting on what you do not want others to do to you.

P1: Make decisions and act regarding things to do to others by reflecting on what you want yourself to do and by observing what others want to do.

The last principle could be reformulated as:

P2: Make decisions and act regarding things to do to others by following what you want to do to yourself and what others want to do to themselves.

With these four principles we have a Confucian CI in distinction from the Kantian CI. As there are four formulations of the Kantian CI that have different connotations, so we have four different principles of the Confucian CI that are together to constitute the essence of the imperative for moral action. The difference between the Kantian approach and the Confucian approach is crystal clear: the Kantian approach makes the moral imperative a formal formula that has little content and provides no easy way of discovery of maxims of action, whereas the Confucian approach makes it possible to discover maxims of moral action in both an easy and simple way because it relates to what is natural for the human self, when the moral will finds itself embodied and situated. There is no incompatibility in the spirit of moral willing by the good will between the Confucian approach and the Kantian approach. There is, however, a fundamental metaphysical difference between the two: the Kantian approach isolates the good will as a transcendent principle opposite to the feelings and desires of nature, whereas the Confucian approach incorporates the good will in the body of feelings and desires so that it finds itself as an immanent principle of inner organization and ordering as well as an integrative principle of leading nature and the body toward a higher level of development within the order of human society.

Based on the four principles of the Confucian CI it is also easy to see how Confucius came to define the four duties regarding the human self and other people in regard to their preservation and development. We may briefly list the virtues to be identified with these four duties implied by the Confucian CI as follows:

1. *Duty to preserve oneself.* To examine oneself reflectively on a daily basis is a duty, cultivating self-respect and making one aware of the deep identity of humanity so that one can care for others. Unlike Kant, Confucius does not explicitly speak against suicide. His statement in *Analects* 15:9, that “there are those who get killed in order to achieve the *ren*; do not seek life at the cost of harming the *ren*”, seems to suggest that it is better to kill oneself rather than doing anything to harm humanity. Oth-

erwise, the main purpose of self-examination is to promote life and develop oneself, so there is no other reason to commit suicide. Poverty and shame are not excuses because a *junzi* would persevere under poverty and shame should lead to self-reform. The *Xiaojing* (*Book of Filial Piety*), reputedly written by Zengzi, says a son with filial piety should not let his body be hurt under normal circumstances and will make all efforts to preserve, otherwise one violates the principle of *ren*. The primary virtues for self-preservation are loyalty (*zhong*, 忠), respect (*jing*, 敬), and sincerity (*cheng*, 诚).

2. *Duty to develop oneself.* The requirement of learning for becoming a *junzi* is a constant and unceasing duty. This is also well-demonstrated in the example of Confucius' ceaseless efforts to learn, to improve, and to self-cultivate. It is a practical necessity that no Confucian could spare for one day. To develop one's virtues also results from one's need to develop oneself so that one will achieve and accomplish a better and a full moral character (*chengren*, 成人). All Confucian virtues that relate oneself to others can be regarded as implying duties to oneself as well as duties to others. The primary virtues for this duty to develop oneself are learning (*xue*, 学), culture and literature (*wen*, 文), moral practice (*xing*, 行), faithfulness (*zhong*, 忠), and trustworthiness (*xin*, 信).

3. *Duty to do justice to others.* The virtue of rightness (*yi*, 义) and the virtue of integrity (*xin*, 信) are the central virtues for doing justice to other individuals and to people in general. Both Confucius and Mencius take these virtues very seriously as they form the cornerstone for the founding of social communities. Specifically, one cannot ignore the importance of family virtues such as filial piety, brotherhood, and even friendliness. In *Analects* 12:7 Confucius says there is no way of dealing with people if there is no trust and honesty.

4. *Duty to care for people.* This has to do with the most central and foundational virtue of *ren*. In our discussion we have seen that it is the ideal end and goal for a person to strive for in both self-development and the development of others. It is therefore both self-regarding and other-regarding because it has to do with humanity in depth, as shared by all people under heaven. As an ideal end for human development and self-development one cannot spare a single moment in such development; for this reason, it is the most necessary duty and cannot be considered a merely a contingent virtue. It is a necessary duty for all occasions because all virtues and all duties are related to this central virtue/duty; as such, this virtue/duty gives unity and vitality to all other virtues and duties. It penetrates the personal duty of respect for life and the fam-

ily ethics of care for parents and siblings and it reaches all people under heaven, all sentient beings, and even all things in the world, as later Confucians have shown. Hence, it is necessary for anyone who is worthy of being developed as a human person and there is no way one can escape from its requirement.

From the above we have exhibited how the Confucian consideration of good will, being embodied and incorporated in the human person, leads to a traceable process of development of virtues and duties. This development also illustrates the unity of the *dao*, as claimed by Confucius. The Confucian system of good will, as embodied in the autonomy of the human person, has clearly come a long way in showing how duties and virtues are rooted in one basic source: *ren* as a deep care for humanity that sustains the whole structure of humanity in time and space. In this sense *ren* must be a perfect duty for anyone to cherish, to sustain, and to advance. We owe ourselves as human persons the care and will for maintaining the present and moving toward the future, and this is why *ren* is a perfect duty in the sense of an ideal end and necessary foundation.

Now I shall point out how this law of morality called *ren* is to be seen as rooted in human nature. It may appear that Confucius did not talk much about the idea of human nature. Zigong says in *Analects* 5:13: “What the Master says about the arts we have heard about, but regarding *tiandao* and nature of man (*xing*) we have heard nothing.” But on the other hand, what he did say has revealed a basic direction for understanding human nature. Confucius says in *Analects* 6:17: “When a man was born, he naturally expresses himself in a just way, but those who do not follow the just way and still survive are lucky to avoid harm and disaster.” There are two key words here: one is *sheng* (生), meaning being born, or giving birth to, and generating; another one is *zhi* (直), meaning straightforward, candid, and just. (Similarly, Wang Baonan in his *Lunyu Zhengyi* (论语正义) quotes Zheng Xuan as saying: “Human nature at its birth is all just.”)

We must point out that there are two aspects of this *zhi*: to speak out or express one’s genuine state; and to express what is genuinely felt regarding the state of affairs and thus to express what one knows or experiences about what is the case and what is not the case. Hence, Confucius’ whole statement is a subtle way of saying that human beings have a natural disposition toward being just, truthful, and honest, implying that there is a human nature in this natural disposition and that this disposition is being just, truthful, and honest. To be

just is to express genuine human feelings with regard to both oneself and the world. People who do not speak out of their heart cannot be just; those who do not respond to the world as they understand it cannot be said to be just either. This shows that Confucius does have a conception of human nature and that he regards human nature as good in the sense that it has qualities of justice, honesty, etc. He would also consider that all persons have this nature or natural inclination. He says in *Analects* 17:2: “all humans have nature very closely the same, their differences come from habit and custom.” Their natures are closely similar in having similar likes and dislikes and having inclinations that could lead to good and from good to bad. It is by habit or practice that people become different by becoming good or bad. The term *xi* (習) has to be understood as a matter of learning and a result of learning and habituation.

Confucius does not speculate on where this natural human disposition comes from. Apparently he conceives it as arising naturally and thus as a natural endowment. But in *Analects* 7:23 he does indicate that his virtue (*de*, 德) is from heaven: “*Tian* (天, heaven) has endowed me with *de* from birth, what could Huantuan do with me?” What is indicated here is that a person could have *de* by a natural endowment from heaven. The so-called *de* is virtue or the capacity to seek eminent performance of good and right in the form of honesty, justice, and benevolence. From this one may infer that human beings have a nature that is endowed by heaven and is virtuous and capable of being realized in honesty, justice, benevolence, and other virtues for maintaining the order and harmony of mankind. They are what the content of morality stands for. This understanding is made explicit in the writings of Confucian disciples such as Zisi and others in the *Zhong Yong* and in the excavated texts recently made available, such as Bamboo Inscriptions in Kuodian, Shanghai Museum, and Tsinghua University in Beijing.<sup>2</sup>

Given this background of Confucius’ understanding of human nature and its content as containing moral dispositions and potency for virtues, one can see that morality as we understand it has a beginning in human nature and it can be naturally expressed and exhibited. But Confucius has gone one step further: he sees his *de* as ultimately *ren* and *ren* as

---

2 Cf. Liang Tao, 郭店竹简与思孟学派 (*Kuodian Bamboo Inscriptions and School of Si-Meng*) (Beijing: People’s University Press, 2008). See also 上海博物馆藏战国楚竹书 (*Shanghai Museum Warring States Bamboo Inscriptions*), 7 vols. (Shanghai: Shanghai Museum Press, 2001–2008). Volumes 1 and 3 are most relevant.

internally experienced and realized by one's will and desire. In the first place Confucius has explained *ren* (仁) as *ren* (人): benevolence as the defining quality of the human being.<sup>3</sup> This explains also how Confucius has referred to *ren* in all major contexts addressing virtues. Besides, he has explicated it, as we shall see, in the unifying principle of the way he comes to understand the truth of humanity: *ren* as the will to achieve good.

*Ren* is a matter of good will in the first place. Confucius has this to say in *Analects* 7:29: "If I desire *ren*, then *ren* has arrived." How could *ren* arrive immediately upon being desired? This is because *ren* is a naturally born quality that can be directly and intuitively grasped. It is within the reach of our natural consciousness as a human being, because we can love others as we will. Besides, to love in the sense of willing good to others (benevolence) and doing good to others (beneficence) are all within our capacity. In addition, it is a matter of free choice that is not hampered by selfish desires and private interests. To desire *ren* is to take the object of the desire as a manifestation of myself, hence it is to experience *ren* immediately as it is part of myself already. But one may also note that *yu* here is the willing of will that is part of the human self. To desire is to will. The immediate experience of *ren* is also a matter of immediate exercise of my will and hence a matter of deciding of my will to become myself. Hence we may express this statement of desiring *ren* as a statement about becoming myself by being myself through my willing:

I desire to be benevolent → I become myself by willing myself to be benevolent → By willing benevolence (an exercise of my will) I become myself as benevolent

The reason we can immediately call for the realization of *ren* in ourselves is that *ren* is the sentiment and will to benevolence or beneficence (i. e., we see benevolence as the basis and drive for beneficence). There is no other way we can do the willing that is the willing of benevolence. Confucius has come to this understanding or discovery through his own reflection on his being himself; he describes it as "to return back to inquire into oneself" (反求诸己) after one has experienced frustrations in reaching one's life goals. It is therefore a matter of deeper experience of self by reflection on the occasions of our experience of the world. This deeper experience of self in actuality is a matter of defining oneself

---

3 See ch. 20 of *Zhong Yong*.



in relation to all our experiences of the world; it requires and gives rise to an integrating motive: *ren*. In later times Mencius simply describes this process as “To return to myself and become sincere (*cheng*) presents the greatest joy. To follow through the care for others is closest to reaching *ren*.”<sup>4</sup> In the *Zhong Yong* one also reads “To be *cheng* is to accomplish oneself (自成).”<sup>5</sup> It further says that “*Cheng* is the beginning and end of things; without *cheng* there will be nothing .” Hence one can see that *cheng* is a quality and action to realize oneself as oneself; it is also a quality that makes things things. “*Cheng* is also the principle which is not for the purpose of accomplishing oneself, it is that by which all things are accomplished”.

One may indeed see *cheng* as a creative act and hence a creative principle whereby one defines oneself and one identifies things as creative products of a cosmic creativity that my creativity is a part of. *Zhong Yong* further suggests: “To accomplish oneself is a matter of *ren*, and to accomplish things in the world is a matter of *zhi*.” This means that benevolence not only defines oneself as a human being but will lend oneself to complete what is to be accomplished by one’s beneficence as one recognizes the inner unity and link between oneself and all things. This of course requires the creative mind of benevolence to acquire knowledge and wisdom (*zhi*). On this ground the Cheng Brothers in the Song Period described *ren* as a power of life-creativity. When we come to the essay on *Ren* (仁说) of Zhu Xi, *ren* is explained as the cosmic virtue (*de*) of creativity, that has a similar power of providing comprehension and life to all virtues.<sup>6</sup>

In light of the above, we can see how *ren* is seen as rooted in human nature and is described as a creative power and principle that preserves humanity as a unity, but also functions as the creativity principle in all things in the world. In this sense *ren* is the cosmic *dao* of the world

---

4 *The Mencius*, 7a:4.

5 *Zhong Yong*, ch. 25 Present and subsequent translations are done by myself.

6 See 朱文公文集卷 67 (*Collected Writings of Zhu Xi*, Part 67) (Taipei: Commerce Press, 1980). It is interesting to note that the word for *de* has been written in the *Chujian* (*Bamboo Inscriptions*) with a double *ren* radical on the left and the word for just or straightforwardness with heart underneath on the right. It may be suggested that the notion of *ren* that has the character of double *ren* on the left side and two on the right side could be construed as two persons sharing the same just mind and hence becomes benevolence or *ren*. The character of *de* hence is described as “to obtain (*de*) within oneself and to obtain it from other people” (内德于己，外得于人， from 徐锴说文解字系传通释).

that generates and sustains the world but is manifested in humanity and human experience of humanity. One may infer that *ren* is actually the incorporation of creativity from the cosmos of heaven and earth that becomes conscious of itself in the human person and thus becomes the defining quality of the humanity in human beings. This amounts to saying that *ren*, as the principle of morality, is rooted in humanity but also rooted or derived from heaven and earth; it shares with heaven and earth the same creative order-giving quality as the cosmic order, as made manifest in Zhu Xi's essay. From this one must also see how the nature of the human person is ultimately founded on *ren* and how *ren* is the ultimate reality of humanity. As to how *ren* can be fully realized and how it is always functionally relevant for providing a source of morality and humanity, the Confucian reply, as we shall see below, is that it is a question of self-cultivation, self-development, and self-realization.

### 3. *Ren* as the Source and Unifying Base of Virtues

Throughout the historical development of the Confucian philosophy of morality and human nature, it is consistently maintained that morality is part of human nature; this nature of ours has been described as a matter of being ordained from heaven in *Zhong Yong*, and morality as virtue is said to be ingrained in one's nature and to need our effort to illuminate and activate it as a power of creative fulfillment and social ordering. To identify nature with morality, as did Mencius and Zisi, has the net result of making us responsible for what we do and what we should do. What we should do is normative: we must see it as a matter of *tianming* in our reflection on our nature. But in the *Yizhuan* this *tianming* has been naturalized to reflect the total creativity of the nature of heaven and earth, as we discover it in our experience. Hence we need not pose an absolute transcendent as the source of the moral principle, as did Kant, but instead point to the natural source in the creativity of heaven and earth as the source of human morality. If the natural source of the creativity of heaven and earth is the principle of ceaseless creativity of creation of life, as is made clear by the naturalizing onto-cosmology of the *Yizhuan*, then it is obvious that the creative source and hence the sustainable foundation of the morality or moral principle must be the very principle of *ren* that Confucius has described in the *Analects* and that has become the basis of human life in the notion of human nature. In other words, human nature is precisely the principle of the morality of creative life

that is embodied in the notion of *ren* and is expressed in the principle of benevolence (*airen*, 爱人) and beneficence (*huimin*, 惠民).

As the ultimate source and foundation, *ren* can be said to be the source and fountainhead of all Confucian virtues, such as *yi*, *li*, *xin*, *zhi*, and even *sheng*. This is because all these virtues have to have their ultimate justification in the power of *ren*. Thus, we can see that *yi* is where *ren* has to introduce fairness and rightness regarding distribution and relations of positions, resources, and opportunities in order to maintain the comprehensive care for a diversity of life in the world. In the process of realizing *yi*, *ren* has sometimes to reach distant goals by extension of practice from proximate centers of human affection as a matter of strategic wisdom. For, unless one is in a position to make equalization a realistic goal, one has to do things from near and extending to the far. The question is whether one forgets the extension and become obsessed with the proximate. *Ren* also leads to *li* insofar as that form of action is essential for making action efficacious as the content of action. Besides, human relationships are as real as what makes relationships possible: the related persons. In this sense *li* is always a matter of how to deliver *ren* in the right way and in the correct form. Just as feeling needs reason for articulation, so is *li* needed for articulating the essence of *ren* or love.

We have seen that reason needs feelings as its manifestation whereas feelings need further reason for articulation and justification. Hence, if we make *ren*, as particularity, derived from the universality of human nature, this particularity needs universality in rationality in order to become rules of moral action governing human relations, rules we may then identify as universality for particularity. Rules of *li* are therefore like maxims of action for Kant: they should embody the universality of practical reason. But as we have argued, Kant forgot to say that this maxim should also embody the principle of *ren*, the particularity of exercising the moral principle in humanity on occasions of life. In this sense moral action has both the form and substance relevant for life, the universality of reason must therefore presuppose the particularity of *ren* in order to apply or relate to actual occasions of life.

Other virtues can be seen as a realization of *ren*, the base virtue: the principle of wisdom (*zhi*) is an aspect of *ren* that makes understanding of self and world available to the activation of *ren*. It is therefore unified with *ren* in such a way that any time there is an awareness of a life situation *ren* should be co-present and any time when *ren* functions as sympathy and empathy there is naturally an awareness that would illuminate

the action in the context of understanding. Both *ren* and *zhi* are complementary with each other so that we cannot speak of one without the other. The virtue of trust (*xin*) likewise goes together with *ren* to ground the ability to trust on the basis of trustworthiness. To trust and be trustworthy is required by the moral principle of *ren* as reflecting the general internal relationships among things and among people. If life comes from one source and we work toward the unity of diversity, how could we proceed without understanding trust and trustworthiness as eminent virtues of human action that results from *ren*?

Finally, we may also mention the virtues of honesty and sincerity: to be true about oneself and to be true about knowledge of things. This is both a principle of *xin* and a principle of *zhi* in action. But it is even more a principle that enables oneself to discover one's deeper self-identity in *ren*. If one is sincere in the sense of bringing out one's true feelings and intentions, one can bring out the deeper identity of *ren* as care and sympathy with others in terms of what one truly cares for and loves in oneself. If one truly care for oneself and loves life, it is not difficult to see how one could care for others and extend love to life in others. In all these cases we see that how *ren*, as the comprehensive principle of morality that consists in comprehensive care and love, is presupposed in all these virtues and how *ren*, with the assistance of practical reason, provides moral justification of all virtues and enables them to become better practiced and also correctly practiced. There is a foundational unity of all virtues derived from *ren* as a source and as an ideal of unity.

#### 4. *Ren* as Both Perfect Virtue and Perfect Duty of Virtue

We have now examined how in the Confucian theory of human nature *ren* becomes the foundational and ultimate source, and hence the fountainhead, of all virtues, providing both unity and a source of vitality to all virtues because it is linked to the very creativity of life in the cosmos of heaven and earth. In this understanding we can make some important observations on *ren* as the principle of morality in contrast with respect for law as the will of morality in Kant. What we are interested in is not their oppositeness but their mutual transformation and internal relationship so that we may make sense of Kantian morality in Confucianism and make sense of Confucian morality in Kant. But to do so requires us to see what is missing in Kant and what poses a problem in Kant. On the other hand, we can also come to see that by accommodating

moral will in Kant in the Confucian theory of nature we will also provide an outlet for introducing the law of morality as an explicit requirement of social order, and to see how morality could be rationalized as rules from individual conscientiousness, so that it may receive a communal form above individual relationships, centered in diverse individuals in a society that has to be organized according to reason and *li*, in both senses of *li*. This is what I call the meeting of minds between Confucius and Kant and the fusion of the horizons of their moral visions.

We must first point out that Kant has incorporated the early Confucian idea of heaven in his reflections on human morality. In his stressing the transcendence of the will that self-legislates for morality, one sees how heaven or God is at work. Confucius has spoken of heaven as what gives him *de* or virtue, so why does Kant refrain from mentioning the transcendent source of universal morality? The answer is that, whereas Kant could be inspired by the idea of the mandate of heaven and even identifies it with God, he comes to see God as absolute transcendence; as (possibly) a Pietist Christian, he fails to see God (cf. heaven) as an all-comprehensive source of life in both past and present and thus as a sustaining creativity of heaven and earth, including human life. In the latter sense heaven is both transcendent and immanent, as the ultimate source and the ceaseless productive force for life. Hence the moral command is the command of creative life and a command for practice of *ren* as care and sympathy for life and hence as benevolence and beneficence toward life.

In cutting off from a deep understanding of the will of heaven, Kant merely sees the moral will as a rational power for imposing the universal law of morality onto human beings without at the same time seeing it as the source of life and care that makes human life universally possible. This makes him unanswerable to the source of life and the fundamental principle of life sustenance that are required for explaining how we have the life form we do. How he comes to this we do not know, but we could suspect that the Hebrew notion of God as lawgiver must have had its impact on Kant's religious notion of God, as much as the Cartesian notion of mind. It is important to see that behind the formal universality requirements of the moral will, there is the ultimate principle of life that makes the universality requirement possible: to make the demand for universality from the moral will is to make life consistent and sustainable as a reality and as a practice. We must see that because of this Confucianism need not insist on heaven as a super-personal persona or absolute will of God, but instead, could naturalize the concept of heav-

en into the concept of heaven and earth, with mankind as a natural outgrowth of heaven and earth.

The naturalization of the transcendent into human nature is a process of the creative formation of human nature as the creative power of self-formation and self-transformation. This means the moral will must become incorporated into the nature of man as the principle of human nature—that is, as what allow human beings to make free choices between good and bad on the one hand and also to allow people to cultivate themselves so that they can make the right choice. Not only this, the moral will also has to be seen to be a power of reflection that considers the creative power of human nature in light of the creative power of heaven and earth, so that one could correct oneself in one's wrong choice in order to be able to fulfill one's life purpose and potentiality. In other words, the moral will is not fixed in one form of command; it is to be seen as the natural powers of free choice, self-cultivation, self-transcendence, self-improvement, and self-perfection. This is the principle of life creativity in the sense of creativity of life without cessation that the *Yizhuan* has attributed to the activities of heaven and earth, nourishing the life of man.

The naturalization and vitalization of the moral will in human beings in the Confucian theory of human nature makes *ren* the perfect moral principle that underlies and is presupposed in all moral principles and moral actions; hence it is considered the arch virtue of all virtues. To call it a virtue is to stress its immanence and its capacity for relative transcendence in the human self. It is to indicate that it has a universal form for humanity but also has an essence of liberating people from their selfish ego and private interests in the creative life itself. Hence this makes Confucius able to say that at 70 he can do anything he wishes without trespassing any moral rules. Morality is his own nature as it is in the ultimate sense and as it is in the original state. But how to make it essential for the daily activities of human beings at any moment of life is the duty of moral cultivation. This is a duty requiring constant care and constant reflection. What is called duty is the sense of mission one must impose on oneself after one realizes how important it is to lead a life of *ren*. Hence Zengzi says in *Analects* 8:7: “A man of education cannot but be persistent and persevering. His mission is heavy and his way is distant. One must see *ren* as one's own mission, is that not heavy? To work hard until one dies, is that not distant?” This is of course a reinforcing response to Confucius' own statement in *Analects* 4:5:

To be rich and have power is what people desire. But if one does not acquire them with the *dao*, one would not stay rich and in power. To be poor and lowly is what people dislike. But if a person does not remove it with *dao*, one would not remove them. Once a *junzi* gets rid of *ren*, how could he be called a *junzi*? A *junzi* will not violate *ren* even during a meal. He must hold it during haste. He must hold it during difficult times.

One can see that as one realizes *ren*, as the ultimate reality and identity of human nature, one must guard it against any depriving forces and one has to see it as one's duty to hold to it so that one will not lose one's true identity. It becomes one's supreme duty to expand the *ren* and apply it to all occasions without exception. Thus, we can come to appreciate and develop all the virtues in different contexts and situations of life, as *ren* will respond to different occasions and situations with proper understanding and proper reflection on oneself. In this sense *ren* provides a motive force to achieve moral wisdom and develop moral vision in one's life practice. Thus a person with *ren* could make a wise judgment as to how to deal with his or her life and care for it, for he could know from his own reflection and knowledge what is the proper way of nurturing himself in the process of self-cultivation.<sup>7</sup> The requirement and urge to cultivate oneself is precisely a requirement and urge from one's self-consciousness and awakened sense of a creative self that will give strength to his moral attitude to develop himself. For acting toward others, to be honest and to be benevolent cannot be and need not be conflicting, because there are occasions when one is immediately called for rather than the other.

In a careful and creative understanding of time one can prioritize one's duties and act them out in a sequence that would harmonize and reconcile any apparent conflicts between them. Hence, there is no necessity to see one duty as more perfect than the other. This is because all duties must be seen in a totality of duties that are all based on the moral will and derived from the ultimate principle of *ren* that substantiates the moral will in terms of the moral sense, moral sentiments, and moral feelings. This is because human nature is an evolving totality of experiences that realize our potentialities for achieving the maximum goodness in our own person relative to different levels of life and relate

---

7 Similarly, the third man in Kant's *Groundwork of the Metaphysics of Morals* (4:423) would have known how to use his talents and develop them in a proper manner without having to conform to rigid rules. A discussion of this argument of Kant's can be found in the sequel, to appear in the *Journal of Chinese Philosophy* (2011).

to different requirements of different communities. It need not be overstressed to say that the very theory of nature that naturalizes the moral will makes it possible for the moral will to manifest itself in terms of human emotions and human desires, not just as human reason divorced from emotions and desires. The central principle of morality is the principle of humanity or *ren*, just as the central principle of humanity is the principle of morality that aims at developing humanity in accordance with the feelings of humanity. In light of this, it is no wonder Confucius maintains in *Analects* 4:4: “If one is devoted to *ren*, there is no moral badness (from oneself).”

In light of what we have said, nothing prevents us from regarding *ren* as a perfect duty, not just as a perfect virtue. It is required and presupposed by all the performances of duties derived from it. Hence, it is not for a moment to be separated and laid on the side. In the same sense that one could become selfish, self-complacent and arrogant toward others in one’s successes and lose sight of *ren* in dealing with oneself and with others, when following what Kant has titled imperfect duties, one could become selfish and obscured in vision in dealing with oneself and others in one’s perfect duties. This of course does not make perfect duties less duties, just as this does not make imperfect duties less duties either. The difference of perfect and imperfect duties hence loses its meaning in a holistic theory of virtues rooted in the theory of human nature with the ultimate principle of *ren* as its basis. This transformation is helpful for Kant, for it enables him to resolve the problem of transcendence of the moral will on the one hand and on the other hand to resolve the problem of justifying the unity of duties as well as resolving potential conflicts of duties.

We may further refer to the transformation of virtues into duties in Kant. This is due to the necessity of conformity with the categorical imperative and enforcement from a source of authority, namely one’s will, that allows one to see the necessity of performing duties. But Kant also compares the moral laws to laws of nature. He recognizes the laws of nature as governing nature as one sees it in Newtonian physics, and thus we wish to see how laws of morality would objectively govern the human world of human actions. But we can see that laws of morality cannot be exactly like laws of nature. We cannot violate laws of nature for any attempt to do so would have consequences that would preserve the observance of the laws. If you fall off a balcony, you will suffer from injury or death as a consequence of gravity. But how about your deviation from and even serious violation of the moral laws of keeping



promises, not lying, and caring for the rights of other people? People who commit those moral crimes often get away and become rich and highly positioned.

When one looks into what has happened in the business world and political arena in the recent history of mankind, one sees how many scoundrels have escaped from prosecution, how many are brought to justice, and how many still enjoy wealth and good life even after a term of imprisonment. From this one sees that moral laws are for ideal types of human beings or for those who have identified with moral laws as their nature and personal identity. Those who choose not to be so identified are on an adventure and a spree of grabbing for profit, luck, and capricious fortunes. They consider themselves brave and smart, but not ethical, and hence there are neither perfect duties nor imperfect duties, from their point of view.

What is then the meaning of speaking of duties and especially of the virtues as duties or duties of virtues for Kant? The sheer purpose for doing so, I believe, is for making those duties laws of legislation in a community or society so that they become enforceable by judicial agencies. This is indeed necessary for some of the moral duties insofar as we need them for guaranteeing maintenance and strengthening the basic forms and structures of social and communal life for the benefits of the majority of people in the long run. In this sense, moral laws have to be eventually conceived as judicial laws of society, for if not so, there cannot be any real meaning in speaking of moral laws. To break promises and to lie are actually now seen as breaking the judicial laws governing contracts. To help others under special circumstances can also be made a judicial law, depending on what usefulness it may achieve. For business and medicine it is apparent that we do need such a stipulation of the duty of benevolence and beneficence and even good will. It is even required for the protection of the innocent against the vicious and calculating. Furthermore, virtues could also be instituted as *li* or rules of proprieties in the sense Xunzi speaks of the *li*.<sup>8</sup> As the last Confucian philosopher in the classical period, Xunzi can be said to have transformed the Confucian virtues into social requirements for main-

---

8 Xunzi sees human nature as morally bad and argues for institutions of ritual (*li*, 禮), rules of action for governing human behavior as practiced by sage-kings of the past. See my article "Xunzi as a Systematic Philosopher : Toward an Organic Unity of Nature, Mind, and Reason", *Journal of Chinese Philosophy* 35.1 (2008), 9–31.

taining an ordered and harmonious society. This was possible because there was a social need to do so. For the individual, this is precisely the work of reason or the *dao* that one can come to in one's reflections on problems of bad behavior that need correction through social interaction and education.

With this said, one may see duty as a result of rational transformation of virtues for a particular purpose and for a particular function required by the common good of the human society and community. In this sense we also see how the principle of *ren* is after all at work in enlisting reasons for maintaining sentiments of humanity that would provide both its formal justification and social authority. In transforming virtues into duties, and hence seeing how virtues could be duties of virtues, one instills the sense of duty for required performance with a certain form, for without such a form we cannot really distinguish virtues from duties of virtues. This form is precisely the form of universality that Kant requires from each moral maxim in its application. With this form we need to look into the content of the maxim and impose the sense of duty to "dutify" the virtues embodied in the maxim. But can we really get rid of the consideration of the content of the maxim in question? For Confucius we need to appeal to one's life sense of sympathy and empathy for the morality of actions. One asks whether what I desire is what others desire and what I do not desire is what others do not desire too. One needs to use one's feelings and imagination in order to make a decision for moral action.

If universalization means that one has to imagine and feel for others, then we have the Confucian golden rules universalized as a duty that makes performance of *ren* a duty of *ren*. But if universalization means simply that one has to rationally accept the result of one's doing something in the formal acclaim of universalizability, then the duty can be anything from connivance to suppression used by a privileged person in power and with privileged positions that would take advantage of the society and community. Hence it is important to see how virtues could be transformed into duties. When virtues become duties they should not lose their status as virtues and should remain at the same time both duties and virtues. They should not lose the link and touch with the individual persons whose moral nature is counted for the ultimate justification of the dutification of virtues. (One cannot really speak of making virtues of empty duties, for they could not be rooted in the nature of individual persons but can be political inventions of the people in power for special purposes in their rule on society.)

Ayn Rand has strongly rejected Kant's notion of duty and called it an anti-concept because it does not answer to anything in human feelings and human nature.<sup>9</sup> However, by introducing the Confucian theory of human nature and the centralizing principle of humanity in this theory, we can save Kant by giving him a proper place in a comprehensive re-consideration of the moral life of mankind, with the ultimate principle of *ren* becoming the perfect virtue for all virtues and also the duty of virtue for all duties of virtues. A new reconfiguration of duties must ensue.

Is benevolence a perfect duty? Yes, it is a perfect duty of virtue because it is a perfect virtue in the first place.

---

9 See Ayn Rand's *Philosophy: Who Needs It* (New York: Bobbs-Merrill, 1982), 95, where she says: "If one were to accept it, the anti-concept 'duty' destroys the concept of reality: an unaccountable, supernatural power takes precedence over facts and dictates one's actions regardless of context or consequences." For her, if we accept the Kantian notion of duty, it would destroy reason, love, values, self-esteem, and even morality itself. That is because, as she sees Kant's position, we have to make duty separate from life and care for life itself.

BOOK ONE:

CRITICAL GROUNDWORK  
FOR CULTIVATING PERSONHOOD

PART I

Epistemological Foundations  
of Personhood



# 1. Self-Cognition in Transcendental Philosophy

*Chong-Fuk Lau*

## 1. Introduction

In the Preface to the First Edition of the *Critique of Pure Reason*, Kant explains his task as follows:

reason should take on anew the most difficult of all its tasks, namely, that of self-knowledge [*Selbsterkenntnis*], and to institute a court of justice, by which reason may secure its rightful claims while dismissing all its groundless pretensions, and this not by mere decrees but according to its own eternal and unchangeable laws; and this court is none other than the *critique of pure reason* itself. (Axi-xii)

According to Kant, traditional metaphysics ended up in a battlefield of endless controversies because it tried to answer questions that transcend every capacity of reason. Kant, therefore, makes it his task to investigate the nature and limits of reason itself, attempting to institute a “court of justice” to settle metaphysical controversies once and for all. The *Critique* is a project of self-investigation and self-examination of reason. This self-reflexive character of Kant’s project is also implied in the concept of *transcendentality*, referring primarily to a peculiar kind of *a priori* cognition: “not every *a priori* cognition must be called transcendental, but only that by means of which we cognize that and how certain representations (intuitions or concepts) are applied entirely *a priori*, or are possible” (A56/B80 f; cf. B25). In other words, transcendental cognition is the sort of *a priori* cognition that deals with the very possibility and conditions of cognition itself. Since these conditions are rooted in our faculty of cognition, “the word transcendental”, for Kant, “never signifies a relation of our cognition to things, but only to the *faculty of cognition* [*Erkenntnisvermögen*]” (4:293).

Kant’s *Critique* is a study of our own faculty of cognition and the principles derived therefrom. Kant offers a systematic division of the mind into different cognitive and non-cognitive faculties; among these the distinction between sensibility and understanding is most crucial to his epistemology. Sensibility and understanding are two stems of

human cognition that are both indispensable and mutually irreducible (A15/B29). While objects are given to us by means of sensibility, they are thought through the faculty of understanding. The former is subject to the forms of space and time, and the latter to the categories. These are the basic tenets of Kant's epistemology. The problem I would like to discuss is the theoretical status of these transcendental cognitions. Are they descriptions about the structure and operations of the human mind? Kant's *Critique* often gives the impression of describing a system of hidden psychological processes or cognitive mechanisms that constructs the world of appearances out of the manifold given from things in themselves. This psychological picture is seriously misguided. I will argue that Kant's discussion of the faculties of sensibility and understanding do not straightforwardly describe the human mind, but the investigation into the faculty of cognition is a conceptual analysis of the structure of cognition of the finite rational being as such.

## 2. Psychological Interpretation

Human beings are equipped with a highly sophisticated cognitive apparatus. We perceive the world through different senses and process the sense-data with an intelligent brain. Does Kant refer to the human cognitive system, when he speaks of the faculty of cognition? How do the faculties of sensibility and understanding differ from, say, the system of the five senses (sight, hearing, smell, taste and touch)? For Kant, the difference is fundamental. The latter refers to the cognitive apparatus of human beings that we know through empirical observation. Descriptions about the five senses are *a posteriori* and contingent. It is totally possible that human beings could have developed more or less than five senses, or different senses than the ones we now have. Nothing rules out the possibility that we someday may acquire the ability to perceive ultrasonic signals, so that we may be able to know what it is like to be a bat. Our possibility in perceiving the world is certainly constrained to a certain extent by the existing five senses, but they are not the *a priori* conditions that account for the possibility of objective cognition. Kant is well aware of the difference (A29/B44). The five senses do not belong properly to the subject of the *Critique*, but rather to that of the *Anthropology*.<sup>1</sup>

---

1 Kant discusses the five senses in some detail in *Anthropology from a Pragmatic Point of View* (7:153 f).

The faculties of sensibility and understanding as well as their spatio-temporal and categorial forms are the *a priori* conditions that the *Critique* has to deal with. Henry Allison calls them “epistemic conditions” in order to distinguish them from merely “psychological conditions.”<sup>2</sup> Human beings are, for example, only capable of perceiving light of a certain wavelength. This certainly belongs to the conditions of human visual perception, but for Allison, “conditions of this sort are not epistemic in the relevant sense, because they have no objective validity or objectivating function.”<sup>3</sup> In other words, these conditions are not indispensable for object formation. Objects appearing under a different color scheme or even without visual properties are conceivable, but the epistemic conditions of spatiotemporal sensibility and categorial understanding are so fundamental that no objects would ever be conceivable if human beings had not developed an appropriate cognitive system incorporating these features. The epistemic conditions, therefore, enjoy a privileged status compared to other psychological conditions.

However, why do space, time, and categories have the indispensable objectivating function, while others do not? The question cannot be answered unless the meaning of Kant’s discussion about our faculty of cognition is correctly understood. The distinction between sensibility and understanding, together with their formal conditions, should not be understood as characteristics that, as Allison suggests, “reflect the structure and operations of the human mind.”<sup>4</sup> Descriptions of sensibility and understanding are essentially different from all empirical descriptions of the human mind. They are not directly about any factual characteristics of a biological species that happens to have developed sophisticated cognitive abilities. All factual descriptions of the human mind are bound to be contingent and known empirically. The conditions of sensibility and understanding themselves cannot be identified through empirical knowledge, including discovery by psychological or introspective observation; otherwise, they would not be *a priori* conditions of experience and objectivity. The human mind can certainly be studied as an empirical cognitive apparatus, but this belongs to the task of cognitive science or psychology and not to Kant’s transcendental philosophy.

---

2 H. E. Allison, *Kant’s Transcendental Idealism: An Interpretation and Defense*, second edition (New Haven: Yale University Press, 2004), 11.

3 Allison, *Kant’s Transcendental Idealism*, 12.

4 Allison, *Kant’s Transcendental Idealism*, 11.



Although Kant does take the “faculty of cognition” to be the “object” of transcendental cognition, it cannot be understood as a cognitive apparatus that functions in a specific way to produce knowledge. The psychological interpretation not only mistakes the faculty of cognition for an empirical object, but more importantly fails to acknowledge that the *conditions* of the possibility of empirical reality themselves cannot be part of empirical reality. The search for the “structure and operations of the human mind” is misguided, if it aims to reveal the empirical workings of the human mind. Every process or operation of the human mind operates and takes place *in time*. Now, if temporality belongs to *a priori* formal conditions of appearance, then the formal conditions that account for the possibility of appearance cannot themselves be temporal processes, nor anything determinable in time. In Kant’s words, “the subject, in which the representation of time originally has its ground, cannot thereby determine its own existence in time” (B422). It does not make sense to explain why and how things must be ordered in a temporal series by describing processes that themselves take place in time. A temporal process cannot be a condition of temporality. The “faculty of cognition” cannot stand for a system of cognitive processes or operations that takes place in the human mind (or the brain) to organize formless data into spatiotemporally and categorially structured appearance. This psychological interpretation mistakes Kant’s project for “a certain *physiology* of the human understanding (by the famous Locke)” (Aix).

### 3. The Nature of Epistemic Conditions

If the epistemic conditions cannot be taken as describing the empirical mind or the phenomenal self, does this mean they refer to something that lies behind the empirical realm of appearance? Are the sensibility-understanding distinction and their formal conditions properties of the noumenal self? Such an interpretation would be in direct conflict with Kant’s criticism of traditional rational psychology’s attempt to prove the substantiality, simplicity, unity and independent existence of the thinking subject from the sole text of “I think.” Kant’s analysis in the Paralogisms unveils the illusion of this metaphysical doctrine, explaining why categories such as substance and causality cannot be legitimately applied to the thinking subject. Kant’s basic epistemological principle rules out the possibility of knowing anything about things in

themselves or noumena. The principle of noumenal ignorance applies to the same extent to the thinking subject itself as to any other external things. It is illegitimate to ascribe the faculties of sensibility and understanding to the noumenal self, since this would presuppose the noumenal self as a substance, with properties attributable to it. It would also make no sense to talk about “operations” or “constructions” of the noumenal self, because temporal and causal categories cannot be applied to it.

According to Kant, no rational self-cognition is possible. I can become an object of cognition to myself only through empirical means such as inner sense. “Yet”, as Kant once conceded, “the human being, who is otherwise acquainted with the whole of nature solely through sense, knows [*erkennt*] himself also through pure apperception” (A546/B574). Using the word “know” or “cognize” [*erkennen*] in this context seems hardly reconcilable with Kant’s basic epistemological tenets. Kant may have chosen a wrong word here, but the concept of “pure apperception” points to another dimension of the problem. Kant does allow a kind of non-empirical access to the faculty of cognition (i. e., a form of intellectual self-consciousness). Although understanding is not a faculty of intuition, it can be “conscious” of the unity of its action “even without sensibility” (B153). The possibility of non-empirical self-consciousness is crucial to Kant’s epistemology. It concerns the pure, original, and transcendental apperception in contrast to the empirical self-consciousness that is in constant flux (A107). Pure apperception refers to “the representation *I think*, which must be able to accompany all others [representations] and which in all consciousness is one and the same, cannot be accompanied by any further representation” (B132). This non-empirical self-consciousness serves as a unifying condition for all representations that can be attributed to a thinking subject.

In Kant’s epistemology, knowledge requires both concept and intuition. As “the consciousness of myself in the representation *I* is no intuition at all, but a merely intellectual representation of the self-activity of a thinking subject” (B278), self-consciousness is still far from being self-cognition (B158). In order to acquire self-cognition, self-intuition is required; this, however, can only be given empirically. Therefore, as argued in the Paralogisms, no rational cognition whatsoever about the thinking subject can be derived from pure apperception. But Kant does admit the possibility of establishing *a priori* cognitions from pure apperception; he calls it “the *transcendental* unity of self-consciousness in order to designate the possibility of a *priori* cognition from it”

(B132). A passage from the *Anthropology* (7:135) further explains the role of pure apperception in logical cognition:

If we consciously represent two acts: inner activity (spontaneity), by means of which a *concept* (a thought) becomes possible, or *reflection*; and receptiveness (receptivity), by means of which a *perception* (*perceptio*), i. e., empirical *intuition*, becomes possible, or apprehension; then consciousness of oneself (*apperceptio*) can be divided into that of reflection and that of *apprehension*. The first is a consciousness of understanding, *pure* apperception; the second a consciousness of inner sense, *empirical* apperception. In this case, the former is falsely named *inner* sense.—In psychology we investigate ourselves according to our ideas of inner sense; in logic, according to what intellectual consciousness suggests.

In order to acquire knowledge of oneself, including one's internal mental state, one has to rely on what is given empirically in inner (and outer) sense. I cannot know anything about myself through pure apperception alone, yet Kant suggests this kind of intellectual consciousness is the source of logical knowledge.

Although Kant's terminology sometimes gives the impression of being psychological, a comparison of his transcendental logic with the general logic will show that his discussion of sensibility and understanding does not refer to mental structures or operations. For Kant, logic is the "science of the necessary laws of the understanding and of reason in general" (9:13; cf. A53/B77). This certainly does not mean pure logic studies how understanding and reason in fact think or operate. There is a branch of logic that "is directed to the rules of the use of the understanding under the subjective empirical conditions that psychology teaches us" (A53/B77), but it is called *applied* logic. Abstracted from all psychological and empirical elements, pure logic is concerned with how understanding and reason *ought to* work (i. e., with the necessary *normative* principles for thought). Logical principles are necessary not in the sense that we cannot think otherwise; instead, their necessity carries normative bindingness (*Verbindlichkeit*). What Kant calls *transcendental logic* also "has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects *a priori*" (A57/B81 f). Just as in the case of pure general logic, transcendental logic does not describe how understanding and reason work, but *prescribes* how their concepts or ideas *ought to* be applied in order to form objective cognition. That is why Kant emphasizes that the transcendental deduction of categories is not primarily concerned with factual questions

(*quid facti*), but rather with questions about entitlement or legitimacy of the use of categories (*quid juris*) (A84/B116).<sup>5</sup>

#### 4. Self-Cognition as Cognition of Cognition

Pure apperception is the consciousness of understanding or spontaneity that enables concepts to be thought. What is revealed in pure apperception is not any internal mental state, but the content of conceptual relations: they are pure and non-empirical because they are not determined in a temporal-causal series. The conceptual-normative principles determined by the spatiotemporal and categorial forms of sensibility and understanding, respectively, can be ascribed to the faculty of cognition, but the latter refers to neither the phenomenal nor the noumenal self. It can be said to belong to the transcendental self or subject, and is

nothing but the simple and in content for itself wholly empty representation *I*, of which one cannot even say that it is a concept, but a mere consciousness that accompanies every concept. Through this *I*, or *He*, or *It* (the thing), which thinks, nothing further is represented than a transcendental subject of thoughts = *x* (A345 f/B404).

The transcendental subject is not a thinking human being that exists in time and processes mental representations with the brain, but only a *theoretical construct* that accounts for the unity and legitimacy of epistemic conditions. The transcendental apperception “I think” is not a factual description of my thinking activity either, but merely a “logical function” (B143, B428). I can certainly use the statement “I think” to describe the empirical fact that I am now thinking, “but the proposition ‘I think,’ insofar as it says only that *I exist thinking*, is not a merely logical function, but rather determines the subject (which is then at the same time an object) in regard to existence, and this cannot take place without inner sense” (B429). The transcendental subject is thus not the *I* that exists in the empirical world and engages in the activity of thinking. It is the ground that makes the indispensable spatiotemporal and categorial structure of objects possible.

---

5 Accordingly, the distinction between understanding and reason in the Transcendental Logic does not primarily address two different cognitive components that human beings in fact possess, but distinguishes between two *types* of normative principles that are valid for making *judgments* and *inferences*, respectively.

Accordingly, the faculties of sensibility and understanding should not be understood primarily as the cognitive components for receiving and processing sense impressions. While understanding provides the normative principles for the application of concepts in judgments, sensibility is responsible for the normative principles that govern the reference to objects. Sensibility and understanding represent two basic and mutually irreducible *functions* for the cognition of objects (i. e., *reference* and *description*).<sup>6</sup> Understanding does not refer to objects directly, but thinks them by applying *concepts* that are nothing but “predicates of possible judgments” (A69/B94), whereas sensibility is responsible for non-conceptual, “immediate” reference to individual objects that predicates can be ascribed to. Jaakko Hintikka suggests “that Kant’s notion of intuition is not very far from what we would call a singular term.”<sup>7</sup> Similarly, Wilfred Sellars understands Kant’s intuition as a function of representing individuals or “*thises*.”<sup>8</sup>

In this connection, space and time are said to be the *a priori* forms of sensibility, not because sensibility receives and processes sense impressions through a spatiotemporal spectacle, but because reference to individual objects has to rely on spatiotemporal coordination. Sellars puts the point as follows:

To intuit is to represent a *this* ... Space and time are ‘forms of intuition,’ not by virtue of being attributes of or relations between things or events in nature, but by virtue of the fact that the logical powers distinctive of ‘this’ representings are specified in terms of concepts pertaining to relative location in space and time.<sup>9</sup>

For Kant, the spatiotemporal framework underlies the possibility of individuation.<sup>10</sup> Space and time form a normative framework that enable

- 
- 6 This is similar to the *demonstrative* and *descriptive conventions* that J. L. Austin takes to be necessary for making meaningful statements. See J. L. Austin, “Truth,” in *Philosophical Papers*, eds. J. O. Urmson and G. J. Warnock (Oxford: Clarendon Press, 1961), 121 f.
- 7 J. Hintikka, “On Kant’s Notion of Intuition (*Anschauung*)”, in *The First Critique: Reflections on Kant’s Critique of Pure Reason*, ed. T. Penellhum and J. J. MacIntosh (Belmont, Ca.: Wadsworth Publishing Company, 1969), 43.
- 8 W. Sellars, “Sensibility and Understanding”, in *Science and Metaphysics: Variations on Kantian Themes* (London: Routledge & Kegan Paul, 1968), 3.
- 9 Sellars, “Some Remarks on Kant’s Theory of Experience”, in *Essays in Philosophy and Its History* (Dordrecht: D. Reidel Publishing Company, 1974), 54.
- 10 Kant does not subscribe to Leibniz’s principle of the identity of indiscernibles. Two drops of water without any inner difference, for instance, can still be held

identifiable references to individual objects to be made. Accordingly, the necessary cooperation of sensibility and understanding in Kant's epistemology can be pinned down to the idea that the possibility of objective cognition presupposes two basic types of normative principles that govern our references to and descriptions of objects, respectively. The point of attributing these forms to the transcendental subject is to express the theoretical status of these forms—that is, their essentially *normative* character. Normative forms are not given from things outside, but are something we *prescribe* to the empirical world. In this sense, “objects must conform to our cognition” (Bxvi). The crucial idea in Kant's Copernican revolution is therefore, as Robert Brandom suggests, “a normative turn.”<sup>11</sup>

Kant's aim is to account for the possibility of objective cognition and reality; this presupposes the epistemic, normative principles discussed previously. If human beings are *in fact* capable of acquiring objective knowledge, our cognitive apparatus must function in a way that complies with the norms. Then, our cognitive apparatus must have sufficiently incorporated the epistemic conditions into its working mechanisms. But whether and to what extent this is the case is a matter-of-fact question that can only be decided by empirical (psychological or cognitive) research. Similarly, whether I am making a valid inference or whether I am thinking correctly according to the laws of logic is a factual matter, while the validity of the logical laws itself is of a normative nature. Insofar as we have reason to believe that human beings do have objective knowledge, we are entitled to assume our cognitive apparatus can be adequately described by the spatiotemporal and categorial forms of sensibility and understanding. But Kant's epistemology is not confined to the human cognitive system. It does not depend on the particular biological, psychological, or cognitive structures of *Homo sapiens*. Kant's project is a conceptual analysis that should be valid to every being that is capable of acquiring objective knowledge, or at least valid to *all finite rational beings*. Kant is relatively conservative at this point, but he does say that “it is also not necessary for us to limit the kind of intuition in space and time to the sensibility of human beings; it may well be that all finite thinking beings must necessarily agree with human be-

---

to be numerically different by virtue of their different spatiotemporal locations (A263/B319 f).

11 R. B. Brandom, *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality* (Cambridge, Ma.: Harvard University Press, 2002), 21.

ings in this regard (though we cannot decide this)” (B72). In regard to the condition of the validity of reason’s ideas, Kant even claims that “given the nature of our (human) cognitive faculty or even the concept that *we can form* of the capacity of a finite rational being in general, we cannot and must not conceive otherwise” (5:401). Kant is concerned with the universal normative conditions, but his analysis seems to have taken a “shortcut”, as if he were straightforwardly describing the structure and operations of the human mind. This is a major reason Kant’s theory often gives the impression of being psychological. If Kant’s investigation into the faculty of cognition provides a sort of *self-cognition*, then it is not cognition of the self, but rather cognition of cognition.

## 2. A Neglected Proposition of Identity

*Robert Greenberg*

[Intuition] takes place only insofar as the object is given to us; but this in turn is possible only if *it* affects the mind in a certain way.<sup>1</sup>

### 1. The Neglect of the Identity Proposition

Kant begins the body of the *Critique of Pure Reason* with the above-quoted proposition of identity; it is part of the foundation of his entire system of knowledge, yet it has been totally neglected in the literature on the *Critique*, at least in the Anglophone portion of it. It would be surprising that it has been discussed in the German or the French literature, or in the non-Anglophone literature in general, for then it probably would have been picked up by English speaking commentators on the *Critique*. So, I am proceeding on the assumption that this fundamental proposition of identity that occurs at the beginning of the *Critique* has not been given any serious treatment by Kant commentators anywhere, if it has even been noticed as having any significance at all.

Aside from some speculation about how this has happened in the literature in general, which I will get to in a moment, I think I have a pretty good idea of how it has happened in the English secondary literature on Kant. It is simply that the English translation of the *Critique* that has been standard until just a few years ago—the Kemp Smith translation—omitted the identity altogether. The Pluhar translation also missed it. Guyer and Wood caught it, but did nothing with it, not even a footnote, nor did either Guyer or Wood discuss it in their own commentary on the *Critique*, as far as I know. If Guyer and Wood caught the identity but did nothing with it, it only makes

---

1 Immanuel Kant, *The Critique of Pure Reason*, trs. Paul Guyer and Allen W. Wood (New York: Cambridge University Press, 1998), A19/B33, my emphasis. Endnotes are stated in the standard manner of Kant scholarship, where “A” and “B” refer to the first and the second edition of the *Critique*, respectively.



more pressing the question of why the identity proposition has not been given any serious attention until very recently, and not publicly at a Kant conference until now. My book, *Real Existence, Ideal Necessity*, seems to be the first thorough treatment of the identity proposition at least in the Anglophone literature, anyhow the first, to my knowledge of the matter.<sup>2</sup>

My speculation as to at least one main reason for the neglect of the identity proposition is that an adequate interpretation of its significance constitutes a challenge to both camps of the widely accepted interpretations of Kant's transcendental idealism—the dual-aspect camp<sup>3</sup> and the double-object camp<sup>4</sup>. Once people are set in their ways, and even have reputations and careers bent on defending those ways, it is hard for them even to *recognize* an obstacle to the correctness of their views. They saw the identity proposition in German, and now in English, thanks to the Guyer-Wood translation, they read it, they understood it, but their own intellectual edifice kept them in a state of denial about its significance. For if they had taken it seriously, they would have had to face the challenge it presented to views for which they had not only become known within the circle of Kant commentary; but more importantly perhaps, the challenge it presented to their opinions of themselves as the duly elected monitors of the holy grail. I must confess that I, too, am afflicted with the same intellectual faintheartedness as affects everyone else. So, in the end we are all in the same defensive posture. The only difference between my posture and those of others is that mine is a challenge to theirs, and I believe it is true. Of course, it goes without saying that theirs is a challenge to still others, and they also believe theirs is true.

I would speculate that a second main reason for the neglect of the identity proposition also has to do with deeply settled philosophical beliefs. But the beliefs I now have in mind are not confined to doctrinaire attitudes belonging to scholars in the two main camps of interpretation of Kant's transcendental idealism. The beliefs are rather widely shared

---

2 Robert Greenberg, *Real Existence, Ideal Necessity: Kant's Compromise, and the Modalities without the Compromise* (Berlin and New York: Walter de Gruyter, 2008).

3 See, for example, Henry E. Allison, *Kant's Transcendental Idealism: An Interpretation and Defense* (New Haven: Yale University Press, 1983), and second edition, 2004.

4 See, for example, P. F. Strawson, *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason* (London: Methuen & Co., 1966).

among epistemologists. They have to do with the opposition in epistemology between causal theories and foundationalist theories of knowledge. This opposition also breaks into two camps, which I will briefly describe in a moment. My second speculation, then, about the reasons for the neglect of Kant's identity proposition, is that commentators on the *Critique* are not averse to joining one camp or the other in this general epistemological controversy, especially if some positions that Kant himself holds suggest taking sides on the issue.

If this second reason for the neglect of the identity proposition is at all valid, a more general, and thus deeper, ground may explain the *first* reason I gave when speculating about the reasons for the neglect of the proposition. Kant scholars may be as susceptible to maintaining the great divide between causal theories and foundationalist theories as other epistemologists are, including Alvin Goldman and H. H. Price, respectively, especially considering some of the things Kant actually said that suggest taking sides on the issue.

## 2. The Grounds for the Identity Proposition

First, Kant clearly does hold to a causal theory of perception: The perceived object plays an essential role in the causal history of the perception.<sup>5</sup> I perceive a particular house because the perception causally originates with the house: The first action involved in my knowledge of the house is that the house affects my senses. The existence of the perception therefore causally depends on the action of the house on my senses. However, Kant also has a foundationalist theory of knowledge.<sup>6</sup> My knowledge of the space where the house exists must refer to the house as it is *given* in my perception of it, if the knowledge is to have objective reality. So, Kant's theory of knowledge is both causal and foundationalist. The same house is causally related to my perceptions of it and, as an object of my perception, instantiates my thought

---

5 For a statement and defense of the causal theory of perception see H. P. Grice, "The Causal Theory of Perception" in *Proceedings of the Aristotelian Society*, Suppl. Vol. 35 (1961), 121–52.

6 For a statement and defense of a foundationalist theory of knowledge see Roderick M. Chisholm, *Theory of Knowledge*, second edition (Englewood Cliffs, N. J.: Prentice-Hall, 1977), 16–33.

about it. Kant's identity proposition is the expression of these two aspects of his theory of knowledge.

### 3. A *Dictum de Omni* and Transcendental Idealism

Kant needs a foundationalist theory to account for objects of knowledge, and he needs a causal theory to account for the existence of these objects. The latter is necessary if he is to distinguish existing objects that instantiate our *a priori* knowledge of nature in general from the abstract objects of our *a priori* mathematical knowledge. Natural objects exist, mathematical objects are abstractions, and Kant needs a theory of knowledge that accounts for the difference. His causal theory of knowledge does the job for him.

The basic proposition of his foundationalist theory is that objects are given to us through intuition and the basic proposition of his causal theory is that representation of the existence of these given objects consists in the effects of the objects on our senses, namely, sensations. Despite the fact that both theories refer to the same objects, each theory has objects of its own. The concept of the identical objects might be considered a primitive concept and the concepts of the objects belonging to each theory might be considered as derived from properties of the identical objects. For example, Strawson considers the concept of a person as a primitive concept that has both psychological and physical properties.<sup>7</sup> Persons are identical objects that have two sets of properties, psychological and physical. These distinct properties can then be the bases of derivative concepts of further objects, only now the objects will have only either psychological or physical properties. That will give us two sets of derivative objects—psychological objects and physical objects. So, we have moved from one set of identical objects with two sets of properties to two sets of derivative objects, each with only one of the two original sets of properties. Let us call these derivative objects “proper objects” of their respective theories. Controversy then ensues over the relation between the two sets of proper objects, whether they are identical, causally related, wholly independent from each other, etc.

The controversy is fueled by a logical *dictum de omni*: Whatever property is predicable of properties of an object is predicable of the ob-

---

7 P. F. Strawson, *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen & Co., 1959), ch. 3.

ject. Joy is an emotion and is predicable of a person, so emotion, being predicable of joy, is predicable of a person. Having weight is determined by gravity and is predicable of a person, so being determined by gravity, being predicable of having weight, is predicable of a person. But of course being an emotion is not predicable of a physical object and being determined by gravity is not predicable of a psychological object.

The lesson of the *dictum de omni* for Kant interpretation is that certain properties that are predicable of properties of appearances or different properties that are predicable of properties of things in themselves, respectively, are all predicable of the identical objects that are referred to by the identity in the proposition under discussion, but are not predicable of their counterpart objects that belong to their counterpart theory. For example, *existence in time*, being a property of appearances, can be *given* and is predicable of the identical objects presently under discussion; so, being able to be *given*, since it is predicable of *existence in time*, is predicable of these *identical objects*. But being able to be *given* is *not* predicable of *things in themselves*. Similarly, *existence apart from sensibility*, being a property of *things in themselves*, is *independent from time* and is predicable of the *identical objects* presently under discussion; so, being predicable of *existence apart from sensibility*, *independence from time* is predicable of the same *identical objects*. But it is *not* predicable of *appearances*. One and the same identical objects both exist in time and are independent of time, depending on whether they are taken as appearances or as things in themselves. The *dictum de omni* is observed in both cases.

Kant's foundationalist theory of knowledge contains the first example concerning appearances, and his causal theory of perception contains the second, concerning things in themselves. But the two theories are brought together in the identical objects that are under discussion. Yet each theory has its own objects—appearances and things in themselves—and they have mutually incompatible properties, and it is about these two sets of proper objects belonging, respectively, to these two distinct types of theories—causal and foundationalist—that the ongoing controversies over the correct interpretation of transcendental idealism are all about.

#### 4. The Objects of a Foundationalist Theory Cannot Play a Role in a Causal Theory of the Existence of the Foundationalist Objects

By way of illustration, I have already said that the proper objects of a physical theory cannot play a role in a psychological theory, and vice versa, though there can be identical objects that have properties of objects that belong to each theory, respectively. I now want to argue that the objects of Kant's foundationalist theory, *appearances*, cannot play a role in his causal theory of the existence of the same objects. So, what I am now going to argue is that appearances cannot be part of the causal explanation of their own existence, although putting it that way may make the point so obvious that argument would seem unnecessary.

The argument is necessary, however, for in *nature*, appearances can be the causes of the existences of other appearances. In fact, in nature, *only* other appearances can be causes of given appearances. The Second Analogy of Experience makes that clear, as have the many commentators who are unhappy with Kant's attempt to employ a non-naturalistic causal explanation of the existence of appearances by employing the concept of things in themselves. But I am going to argue nonetheless that that is exactly what Kant does. The saving grace of my interpretation of Kant's causal theory of knowledge, if it has one, is that the set of identical objects—the ones that are presently under discussion—have properties not just of the proper objects of causal theories of knowledge, but also of the proper objects of foundationalist theories. Thus, while the identical objects have properties that figure in the causal history of empirical objects that are given to us, they also can have properties that belong to these very empirical objects whose existence their causal properties help explain.

Appearances cannot play the assigned role in Kant's causal theory for several reasons. The reason I would like to present here trades on Frege's theory of the informativeness of identity propositions.<sup>8</sup> One might at

---

8 Gottlob Frege, "On Sense and Reference," tr. Max Black, in Peter Geach and Max Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Basil Blackwell, 1952). For a complete development of this Fregean argument as well as other reasons that appearances cannot play a role in Kant's causal theory of the existence of the foundations of knowledge, see my *Real Existence, Ideal Necessity*, ch. 2, section 7.

least have to tolerate, if not accept, Frege's theory if one is going to be at all happy with my argument. The identity proposition that has been under discussion in this essay is not an analytic proposition, for Kant does not understand "affects the mind" as "given to us," and conversely, he does not understand "given to us" as "affects the mind." Otherwise, things in themselves could be given to us and mathematical objects could affect the mind, both of which are impossible according to Kant's understanding of the expressions. Accordingly, things in themselves can affect the mind and mathematical objects can be given to us. As undetermined objects of empirical intuition, appearances are given to us, and for Kant it is analytic that appearances are such objects—that is how, *inter alia*, he thinks of them.

Now, how are we to understand the objects that affect the mind? The proposal under examination is that the objects can be understood as appearances. This means that for Kant the sense of "affects the mind" would be the sense of "appearance". But the sense of "appearance" contains the sense of "given to us," or something to that effect. According to the proposal, therefore, the identity proposition would be analytic! But the major premise of the argument is that the identity proposition is *not* analytic. The conclusion of the argument, therefore, is that the proposal is false. It should therefore be rejected.

This is a simple argument that trades only on a tolerance of Frege's analysis of the informativeness of identity propositions—namely, that they cannot be analytic, as understood in terms of his distinction between sense and reference—and Kant's uses of some of his basic terms. Let it not be supposed that one might object that it can be informative to learn a sense associated with a given expression. We are assuming our already having learned the sense of an expression and going on from there. The expression as a physical shape is not an object of our discussion. Indeed, we are talking about in what senses Kant uses given expressions, not about the relation between an expression as a physical shape and a sense.

## 5. The Objects of a Causal Theory Cannot Play a Role in a Foundationalist Theory, Either

Obviously, things in themselves cannot be part of any foundation of knowledge for Kant, since they can only be thought and cannot be given. But cannot the identical objects of the proposition under discussion play a role in a causal theory? The problem with the suggestion is probably already obvious. As soon as the identical objects are identified with one of the proper, counterpart objects—either appearance or thing in itself—they lose their identity as the identical objects that have two distinct sets of properties. It is essential to them that they are rather *common objects* of both sets of properties. The alternative is: either one set of objects with two distinct sets of properties, or two distinct sets of objects with one set of properties. The first alternative is the dual-aspect interpretation of transcendental idealism, and the second is the double-object interpretation. While my interpretation is dual-aspect, I have also tried to account for the double-object one. The attempt has been made by highlighting a certain contrast between causal and foundationalist theories of knowledge. The basic idea has been that the double-object interpretation can be derived from the dual-aspect interpretation.

## 6. A New Interpretation of Transcendental Idealism

Transcendental idealism requires that these identical objects are thought independently of the relation in which they can stand to sensibility. Standing in that relation, they are appearances. The concept of an appearance is a *relational* concept: it holds for these independent objects just in case their relation to sensibility is *affective*. As such, they are represented as objects of Kant's foundationalist theory of knowledge, that is, as *appearances*. On the other hand, they can also be thought as being objects of the understanding *alone*, or as objects of reason. This is the counterpart property that can be ascribed to these independent objects, and now they are represented as objects of Kant's causal theory of knowledge. Just as psychological objects cannot be physical objects, and conversely—to refer back to our earlier analogy for a moment—so objects of one theory cannot be objects of the other.

Identical objects can have two properties: one is a relational property, that is, as objects that can stand in relation to sensibility; the

other is a non-relational property, that is, as being independent of sensibility. This explains why one and the same objects can have incompatible properties at the same time. It is because the two types of property, at least in this case, do not interfere with each other. Of course, one has to keep one's thoughts about them in these two respects distinct or one would be confused. The independent objects, of course, must be thought independently of either property. Yet this entails at least *thinking* them, and that means they are being thought through concepts of the *understanding*. As represented by means of the understanding, they are thought *independently of sensibility*. This is to think their independence from sensibility *adverbially*, as a *manner of thinking* them. But to think them in this manner, that is adverbially, is *not* thereby to think of them *as being thought independently of sensibility*, or more simply, it is not thereby to think of them as *being independent of sensibility*. The latter is to ascribe a property to them. As objects of such ascription, the identical objects are *things in themselves*. Represented by means of the understanding, however, they can also be *determined* by sensibility, as is done in the Aesthetic and Analytic portions of the first *Critique*. Therefore, the identical objects can be represented by means of the understanding and then they can be *determined* both by sensibility, as objects of empirical intuition (i. e., as appearances), and by the understanding, as objects of the understanding alone (i. e., as being objects of reason that are independent of sensibility), as is done throughout the first *Critique*, but especially in the Dialectic.

The more specific concern of this essay, however, has not been this new way of interpreting transcendental idealism, but rather the idea that Kant's theory of knowledge consists of two theories of knowledge, one causal, the other foundationalist, and unless both are given their due, not only will Kant scholars be continually vexed by interminable problems of interpretation, but Kant's original idea that epistemology should not try to get on without both types of knowledge theory will be lost on those non-Kantian philosophers who are currently working in epistemology<sup>9</sup>.

---

9 Strawson, the Kantian, is a notable exception among epistemologists concerning the relation between causal and foundationalist theories of knowledge, as he, like Kant, combines the two theories in a single, unified theory of knowledge. See his "Perception and Its Objects", in *Perception and Identity: Essays Presented to A. J. Ayer*, ed. G. F. Macdonald (Ithaca, NY: Cornell University Press, 1979).



### 3. Kant and the Reality of Time

*Kwok-Kui Wong*

#### 1. Introduction

In Kant's *Critique of Pure Reason* the major problem confronting the unity of personhood is that time is a flux; according to Kant, this is the basic feature of time as inner sense. He says in many instances that nothing abides in the flux (A364, A382), not even the "I". So he says, "No fixed and abiding self can present itself in this flux of inner appearances" (A107) This problem is however not only a challenge for personhood, but also for the very possibility of experience as such. Concerning time as succession Kant says thus (A183–4/B226–7): "For in bare succession existence is always vanishing and recommencing, and never has the least magnitude." If nothing abides and nothing has magnitude, how can we have any possible experience? This ancient problem of "everything flows" therefore also has its presence in Kant. So he says in the Third Paralogism (A364): "Although the dictum of certain ancient schools, that everything in the world is *in a flux* and nothing is *permanent* and abiding, cannot be reconciled with the admission of substances, it is not refuted by the unity of self-consciousness." The fact that Kant has looked for the basis of the unity of consciousness in transcendental apperception rather than the empirical consciousness does not mean that the problem of flux has been resolved. We may summarize the question in this way: if time is a flux, then how can there be anything abiding in this flux so that we may have any form of a unity, whether it is a unity of an object of experience, unity of consciousness, or even of a person?

One strategy of considering this question is to confront Kant with a classical exposition of the question about the reality of time, namely the *Confessiones* of Augustine (354–430). For Augustine's doubt about the ontological reality of "now" is another formulation of the notion that "nothing abides" in the temporal flux. Augustine reasons: if time is essentially made up of the past, present, and future, the past cannot *be* because it has passed, the future cannot *be* because it has not yet come,

while now cannot *be*, for if it can, it cannot flow and pass away, and therefore will not be time but eternity. If the existence of now is founded on the fact that it always becomes *not-being*, then now itself cannot have *being*. So “now” can only be an infinitely small division between the past and the future. Therefore, the core of the ontological problem of time is the extension of now, or an instance. In other words, Augustine’s question is: how can *now* have extension and magnitude even though no part of it will become either the past or the future, thus seeming to render it non-existent?

Kant’s discussion of time in the first *Critique* is of course not a direct response to Augustine’s question. One of the many significant differences between Augustine and Kant is that the former’s treatment and answers to his own question, namely *attentio* and *distentio animi*, remain in the framework of pure time (i. e., time without reference to outer spatial experience). Kant, on the other hand, makes a conscious distinction between time as inner sense and time in relation to objective experience. The aim of this essay is therefore to examine how far Kant’s treatment of time, not merely as purely subjective “inner sense”, but also as part of the spatial-temporal experience of the objective world, can offer a different answer from that given by Augustine himself.

## 2. The Transcendental Aesthetic

To begin with, we should first examine the Transcendental Aesthetic, for this is where Kant treats time thematically. Here he outlines the basic features of time: time has only one dimension, different times are different parts of the same time, while the infinite extension of time can only be imagined with the appearance of objects of experience. Kant thus concludes: first, we cannot suppose the objective existence of time after abstracting all appearances; second, time itself may appear to us as inner sense; third, time is the formal condition of all appearance, a pure form of intuition. Therefore, Kant concludes that time possesses a certain ideality rather than objective reality.

Can Kant’s conclusion answer Augustine’s question? The answer is yes and no. We may conclude that Kant has agreed to Augustine’s doubt that time cannot have objective existence. On the other hand, time as inner sense may flow *ad infinitum* even without corresponding external appearance. Therefore, we may only imagine the form of time with the appearance of objects of experience as a line leading to infinity, while

without appearance there can be no objectively existing time. The ontological existence of time may not be founded on the fact that it is a pure form of intuition. It only shows that whenever we intuit an object of experience, it must be intuition in such a form.

Therefore Kant has not hinted how to answer the question concerning the extension of time in the Transcendental Aesthetic. The reason may be that here Kant is only examining space and time as basic forms of sensibility, and he has not yet moved to investigate experience with objects of appearance. The time Kant is talking about here is only a form for all possible experience, and therefore it must be a single form, one time rather than a period of time. If we are to found the extension of time on a period of time or a certain duration in the flow of time, then beyond this period of time, there must be another longer period of time until we must imagine an ideal time that extends *ad infinitum* as the condition of experience. This would again turn into a form of experience, a pure form rather than a real entity, and this differs significantly from what Augustine is looking for. Time as a condition of experience only indicates that when an object of experience appears, it must appear “in time” (i. e., a point in the infinitely extended line with a position after the past or before the future). It does not prove that the past and the future really *are*.

### 3. The Analogies of Experience

Apart from the Transcendental Aesthetic, time also plays an important role in schematism, and so also in the Analogies of Experience. The three analogies are each related to time: permanence, succession, and co-existence. We will mainly look at the permanence of substance, as the other two analogies may not provide an answer to our question about the extension of time. Succession and co-existence cannot prove the extension of time. Succession alone without permanent substance is only alteration, a series of unextended points, while co-existence without succession is the denial of the flow of time. However, this extension of substance involves many ontological contradictions: on the one hand, it permeates through all time, occupying the past, present, and the future. If so, this substance should have extension in time. However, Kant also says (A183–4/B226–7):

If we ascribe succession to time itself, we must think yet another time, in which the sequence would be possible. Only through the permanent does existence in different parts of the time-series acquire a magnitude which can be entitled duration. For in bare succession existence is always vanishing and recommencing, and never has the least magnitude.<sup>1</sup>

Thus Kant asks again whether this substance has a beginning and an end. In order for the perception of change to be possible, there must be something that does not change and underlies the changing appearance so that change will not become mere alteration. For change is essentially the alteration of accidents accompanying the unchanging substance. In this process the appearance and disappearance of a certain state or condition must have a definite point of time, and before this point of time there must be another point of time that we can perceive. Otherwise this point of time may not become a part of our possible experience as empty time cannot be perceived.

This characteristic of substance has provided an absolute basis for extension. This extension into infinity is absolute because we cannot conceive of change without imagining a substratum underlying this change. Therefore, this substance goes from one extreme of minute instance without extension to the other extreme of infinite extension. However, we cannot be satisfied with this characteristic of substance as an answer to the question of the extension of time. The first reason is that, as an analogy of experience to synthesize the manifold of appearance, Kant has not established its objective existence. Following the postulate of Newtonian physics, Kant posits a permanent substance in order not to violate the principle of “*gigni de nihilo nihil, in nihilum nil posse reverti*” (3:223).<sup>2</sup> However, Kant also points out that this substance is only an assumption, but cannot be stipulated dogmatically as objectively existing. Second, if this substance has neither beginning nor end, but permeates all experience, we will ask the question whether this permeating existence and the succession of instants are two parallel lines running along each other without intersection. Is it merely the opposite of a succession of instants without extension, an assumption that connects all these infinite number of instances? How can the permanence be connected with the instances, so that the latter can have extension as a real entity?

---

1 Immanuel Kant, *Critique of Pure Reason*, tr. Norman Kemp Smith (New York: Modern Library, 1968).

2 “Aus nichts werde nichts hervorgebracht, in nichts könne nichts zurückkehren.”

How can this entity have existence in time with duration? All these questions do not find satisfactory answer in the First Analogy, as Kant has not explained how this substance can become an existence in time with duration through conceptual determination. Therefore, given its metaphysical nature, we cannot find any proof of the extension of time in this substance.

#### 4. Kant's Transcendental Deduction

Now we look back at one of the most important chapters in Kant's *Critique*, namely Transcendental Deduction. In the A-Deduction Kant speaks of a concept produced completely *a priori* (A95), which does not come from real experience, but can be applied to objects of experience. Kant says that if there is such a concept, it must be the condition of all possible experience, and therefore transcends all objects of experience. Further, Kant calls this concept a "pure concept of understanding".

Further Kant says that in order for experience to be possible, or in other words, in order to make the manifold into appearance for knowledge, we must perform a certain synopsis: two things, receptivity and spontaneity, must work with each other. It is well known that when Kant talks about spontaneity, imagination is usually referred to. There is then the distinction between "productive imagination" and "reproductive imagination", and the former is spontaneous. From these two points we may conclude that the constitution of the unitary object of experience requires the participation of the receptivity of sensibility and the spontaneity of the subject (including productive imagination). We may ask therefore how this unit is to be constituted. How does spontaneity work with receptivity so that the former does not become fantasy but should have certain objectivity? And the most important question is: what role does time play here?

In the first step of the A-Deduction, namely "On the Synthesis of Apprehension in Intuition", Kant emphasizes the importance of time, and indicates that time is the basic form for all knowledge (A98–9). We may even say that before knowledge begins, when our intuition of the world has contact with the world of experience, time has become involved. Kant is of course not directly addressing the Augustinian question about the reality of time in this part. Yet we may see in the text that if Kant says that intuition takes place in an instance, in order for intu-

ition to be able to constitute knowledge, this raises the question of how this instance can engage in the flow of time.

In his interpretation of Kant, Heidegger argues that in Kant's arguments regarding transcendental synthesis and transcendental deduction, imagination is the basis of the reality of time, especially the so-called "transcendental imagination". Heidegger's interpretation of course induces much criticism from Kantian scholars. However, he has given valuable hints in the understanding of the role of time. Heidegger says this about the relation between intuition and an instance:

... for every now is now already just arrived [*schon soeben*]. In order for the synthesis of apprehension to give the current look perfectly in an image, it must be able to retain as such the present manifold which it runs through; and at the same time it must be pure synthesis of reproduction.<sup>3</sup>

To put it simply, intuition is not passive receptivity, but at the moment of intuition it has already been involved in the constitution of impressions. Otherwise the constitution of knowledge has nowhere to begin. For in order for knowledge to be possible, the faculty of knowledge of human beings, whose existence is also in time, must also have the flow of time as its basis. The so-called "transcendental imagination" turns an instant impression that in principle has no extension in time into something "temporal". This is the basic faculty of human consciousness but also at the same time constitutes our finitude. They are two sides of the same coin, and we cannot ask whether the same is true in the real world.

Now we may look back at Kant's explanation of transcendental imagination to see how it works. Imagination is understood by Kant as the representation of an object even in the absence of it. According to Kant, it uses what we have seen and observed as raw material for further imagination, and it cannot function independently of this raw material. The so-called "reproductive imagination", rather than being pure fantasy, must be based on certain "empirical laws". For example, when we see a car travelling by, and then our vision of the car is blocked by another vehicle, we can still project the time when the car reappears from the back of this vehicle because we can estimate the speed of this car based on the first impression of it. Kant argues that apart from "reproductive imagination", there is also "productive imagination". This very concept involves a contradiction with itself because imagina-

---

3 Martin Heidegger, *Kant and the Problem of Metaphysics*, tr. R. Taft (Bloomington & Indianapolis: Indiana University Press, 1997), 128.

tion is defined as the projection of the further activity of an object of experience without its presence. However, if imagination is productive rather than reproductive, does this mean that it is not based on any first experience and the related empirical laws? Kant says this in B-Deduction (B151):

But inasmuch as its synthesis is an expression of spontaneity, which is determinative [*bestimmend*] and not, like sense, determinable [*bestimmbar*] merely, and which is therefore able to determine sense a priori in respect of its form in accordance with the unity of apperception, imagination is to that extent a faculty which determines the sensibility a priori; and its synthesis of intuitions, conforming as it does to the categories, must be the transcendental synthesis of imagination. This synthesis is an action of the understanding on the sensibility; and is its first application—and thereby the ground of all its other applications—to the objects of our possible intuition. As figurative, it is distinguished from the intellectual synthesis, which is carried out by the understanding alone, without the aid of the imagination. In so far as imagination is spontaneity, I sometimes also entitle it the productive imagination, to distinguish it from the reproductive imagination, whose synthesis is entirely subject to empirical laws, the laws, namely, of association, and which therefore contributes nothing to the explanation of the possibility of a priori knowledge. The reproductive synthesis falls within the domain, not of transcendental philosophy, but of psychology.

We can see from the above passage that the so-called “productive imagination” is not determined by external sensibility but is rather determinative. How does this imagination come about? Heidegger argues that the imagination Kant talks about is not pure *creatio ex nihilo*, but corresponds to the conditions of experience. It is original and receptive at the same time.

We may explain this by analyzing the three steps in Kant’s A-Deduction. First, the so-called Synthesis of Apprehension in Intuition is perhaps the most important one, so it deserves detailed examination (A98–9):

Every intuition contains in itself a manifold which can be represented as a manifold only in so far as the mind distinguishes the time in the sequence of one impression upon another; for each representation, in so far as it is contained in a single moment, can never be anything but absolute unity. In order that unity of intuition may arise out of this manifold (as is required in the representation of space) it must first be run through, and held together. This act I name the synthesis of apprehension ...

Kant seems to point out that for intuition in the manifold to be possible, time must somehow be involved. However, the words Kant uses, “run

through” (*Durchlaufen*) and “held together” (*Zusammennehmen*), seem to imply an activity in time. Despite the extreme brevity of this instance of intuition, a certain time seems to be already involved, otherwise this process “running through” and “holding together” may not be possible.

We may analyze this situation in the following way. For example, when we see a manifold, in the process of “running through” and “holding together”, time is a hidden transcendental condition because in this process, even in the instance of time without extension, we must suppose that a certain passage of time has taken place. For if we look back at this intuition after a period of time and trace back the process of how this flux of time becomes a meaningful appearance making knowledge of the outside world possible, then every instance has a tendency to move over from its infinitely small instance and become part of this meaningful memory. To put it in terms of what Schelling says in his *System of Transcendental Idealism*, philosophy is a “free imitation, free recapitulation of the original series of actions”.<sup>4</sup> It means roughly that when we look back from the entirety of a piece of epistemologically meaningful experience, then each infinitely small instance that makes up such an experience should also contain within itself a piece of this entirety, whilst this entirety is only then an obscure background and can only enter into our consciousness through multiple repetition. This is the line of argument often used by German Idealism from Schelling to Hegel, who obviously inherited the legacy of Kant’s transcendental deduction. The difference is only that Kant wants to argue the other way round: to deduce the reality of the entirety of experience from the possibility of individual experience. However, since our purpose is not the same as Kant’s transcendental deduction, we may follow a line similar to that of the Idealists (i. e., to prove the reality of time from the entirety of experience).

How does this argument proceed in concrete terms? To answer this question, we may further examine to Kant’s distinction between “productive imagination” and “reproductive imagination”. In experience, sensibility and imagination about individual objects must on the one hand abide by the empirical laws. Yet the possibility of the unfolding of these empirical laws relies on the other hand on certain transcendental laws. Heidegger refers to it as a “holding open of the horizon” (“*das*

---

4 F. W. J. Schelling, *System of Transcendental Idealism*, tr. Peter Heath (Charlottesville & London: University Press of Virginia, 2001), 49.



*Offenhalten des Horizontes*”)<sup>5</sup>, a “space of play” (*Spielraum*)<sup>6</sup>, or a “pure picture” (“*reines Bild*”). This open space is however not derived of any content. Only that this content cannot appear by itself, but must be brought about by the presence of a certain object and its related empirical laws in this space. Its most basic form is the features of time as discussed by Kant in his *Metaphysical Exposition of the Concept of Time*: for example, that time must be in constant flux in a single line. We may pick up two most basic features for our discussion. The first is the contrast between co-existence and succession (A30/B46). Kant argues that only in temporal experience can these two basic forms meet each other (*antreffen*) (A32/B48). Second, Kant also points out that time is the necessary representation of all possible experience, that time must accompany all experience.

According to the two points above, spatial-temporal experience must have the involvement of co-existence and succession. We cannot conceive pure co-existence in our actual experience without succession, for it means that time has stopped flowing. By the same token, we also cannot conceive pure succession without co-existence, for otherwise experience would become a pure flux: we could not tell what is what and could not make any distinction between different objects, and space would have no role to play. The meeting of co-existence and succession must take place by the formation of a unitary object of experience that permeates in time. This object, however, taken in itself, has no particular shape or form, as we cannot imagine a pure object, as much as we cannot imagine a pure triangle without specifying the size of each angle.

Therefore, in any spatial-temporal experience, in principle, we have to decide in the smallest instance of time which part of the manifold belongs to succession and which part belongs to co-existence. At the same time, we also have to decide whether another object co-exists with the object in the same space. However, we must emphasize that this decision cannot be made arbitrarily, nor is it merely determined by our consciousness, but must be done according to our impression of this object and its related empirical laws. To use the example we have discussed earlier: from the empirical laws we can imagine when the car will appear again after the brief period of its absence because we can use the first impression of this car and project its speed. However, behind this

---

5 Heidegger, 90.

6 Heidegger, 59.

imagination another imagination is going on every moment: the picture (*Bild*) where succession and co-existence meet each other. This vehicle as a unit that permeates in time, the part that changes in the flux of time, and the other vehicle that blocks the first one—all of them are permanent units in time. A preliminary constitution of unitary objects of experience must have taken place in a certain way, though this constitution may not be clear in the first stage of “running through” and “holding together”. What is certain is that in the first impression succession and co-existence must participate, otherwise any permanent unit as the basis of further imagination may not be possible.

In the second step of the A-Deduction, the Synthesis of Reproduction in Imagination, the participation of succession and co-existence is even clearer. When we imagine a vehicle appearing from behind another vehicle, this permanent unit is not simply an agglomeration of the manifold, but we may also perceive its form of permanence and its difference from other permanent units. For example, we may tell that its color and shape may not change in time, while its position in space may change. The same applies to the other objects that make up the background of this movement and may not change. In both cases, succession and co-existence as two forms of time have been involved.

In the third stage, the Synthesis of Recognition in Concept, the vehicle has appeared again. While we recognize that it is the same vehicle, this recognition assumes a knowledge of the essence of this object of experience so that this recognition is possible. We recognize this object not only according to its shape and color, but we also that assume there is a causal relation between its shape and color in its first impression and the speed of its travel, though this relation is actually not established. Kant calls this transcendental object “unknown = X”. At the same time, Kant also says that the constitution of a concept is the key to this recognition. In this way, this unitary object is not only an assumption of a permanent object, but something that must be present in all possible external experience in order for knowledge and judgment to be possible.

## 5. Conclusion

To go back to Augustine’s question, we may say that for Kant, a pure instance without extension does not exist because succession and simultaneity must participate as part of our possible experience that forms the

entirety of our spatial-temporal experience. To question the reality of an instance is to question the very possibility of spatial-temporal experience. The answer to the question of the extension of time cannot be given from time alone, but must be related to the constitution of a unit of knowledge.

This essay constitutes only a small attempt to answer the question concerning the extension of time vis-à-vis time as pure flux, and it is an even smaller attempt as far as the unity and personhood is concerned. However, this attempt is undertaken with the belief that any answer to the question about the unity of personhood must confront the problem that “nothing abides” in the flux of time, rather than evading it by seeking refuge in the unity of transcendental apperception. The involvement and convergence of time and space in the constitution of a unitary object of experience is, therefore, the first step toward exorcising the ghost of the heraclitean flux in Kant’s first *Critique*.

## 4. The Active Role of the Self in Kant's First Analogy

Gregg Osborne

### 1. Introduction

Kant gives two versions of the principle to be proven in the First Analogy. "All appearances contain that which persists (**substance**) as the object itself," he announces in the first edition of his *Critique of Pure Reason*, "and that which can change as its mere determination."<sup>1</sup> "In all change of appearances substance persists," he proclaims in the second, "and its quantum is neither increased nor diminished in nature" (B224). In spite of intensive study and debate over the course of several hundred years, there is still no consensus on the nature, or even the number, of the proof (or proofs) he then goes on to offer.<sup>2</sup> My own conviction is that there is only one basic proof, that crucial parts of it are not spelled out in the text, and that a most essential claim is somewhat misstated in the first edition.<sup>3</sup> Given the length limitations of this essay, however, the full grounds of this conviction cannot be spelled out

---

1 A182. Citations from and references to Kant's *Critique of Pure Reason* follow the standard practice of referring to the page numbers of the first and/or second editions, "A" to the first and "B" to the second. For passages that appear in both editions, the relevant page numbers in both are indicated.

2 Such commentators as Norman Kemp Smith, A. C. Ewing, Robert Paul Wolff, Paul Guyer, and P. F. Strawson all take there to be several. Others, such as Béatrice Longuenesse, explicitly affirm that there is only one. Opinions on the nature of the proof or proofs vary too widely for brief summary.

3 "Our **apprehension** of the manifold is always successive, and is therefore always changing," begins Kant in that edition. "We can therefore never determine from this alone whether this manifold, as object of experience, is simultaneous or successive, if something does not ground it **which always exists**, i. e., something **lasting** and **persisting** ..." (A182/B225) What he should actually say in the second sentence, in my view, is that we can never determine from "this" alone whether the changes (i. e., cases of coming to exist and ceasing to exist) that constantly take place in apprehension are objective or take place "in the object" if something of the sort he describes does not ground the manifold.

here. What I will try to do instead is explain the basic structure of the proof I take Kant to have in mind all along but fail to express with full clarity in either edition. At the very basis of that proof, I will suggest, is a revolutionary account of the role of *activity* on the part of the self in the generation of one of the most basic and ubiquitous features of what is commonly and naturally referred to as experience. The feature in question is our putative awareness of coming to exist and ceasing to exist that is objective as opposed to merely subjective in nature. By emphasizing this revolutionary account in a mildly technical analysis of the First Analogy, I hope to contribute to our collective exploration of Kant's conception of human personhood.

## 2. Kant's Response to a Very Basic Question

A very basic question that concerns Kant in the First Analogy is how we can perceive that something has come to exist or ceased to exist. His concern with this question can be seen in the fourth sentence of the proof (or version of the proof) added to the second edition and is also front and center in a striking passage from A188/B231. What Kant most clearly implies in the first of these passages (B225) is that there must be a substratum in the objects of perception that does not come to exist or cease to exist and that in relation to it all cases of coming to exist and ceasing to exist can alone be perceived.<sup>4</sup> What he implies with equal clarity in the second is that a case of coming to exist or ceasing to exist cannot be perceived as such "unless it concerns merely a determination of that which persists."<sup>5</sup> The phrase "that which persists" in this second passage must refer to the substratum in the objects of perception that he refers to in the first. What Kant clearly implies in the combination of these passages, therefore, is that a case of coming to exist or

---

4 "... it is in the objects of perception ... that the substratum must be encountered that represents time in general and in which all change ... can be perceived in apprehension through the relation of the appearances to it." In order to represent time in general, as the second sentence of the relevant proof (or version of the proof) clearly indicates, the substratum in question must last and not "change". To *change*, it can be seen a few pages later at A187/B230-1, is not merely to undergo alteration but rather to come to exist or cease to exist.

5 "... arising or perishing ... cannot be a possible perception unless it concerns merely a determination of that which persists ..."

ceasing to exist cannot be perceived as such unless (a) there is a substratum in the objects of perception that does not come to exist or cease to exist, and (b) the relevant case of coming to exist or ceasing to exist concerns merely a determination of that substratum. A natural way to parse this, in my view, is that we can be aware, on the basis of apprehension, that something has come to exist or ceased to exist only if its doing so was a mere change in the determinations of something that already existed and continues to exist. Let us call this claim  $\beta$  and ascribe it to Kant:

$\beta$ : We can be aware, on the basis of apprehension, that something has come to exist or ceased to exist only if its doing so was a mere change in the determinations of something that already existed and continues to exist.

We are now faced with at least two major questions. First, why would Kant subscribe to  $\beta$ ? And second, why would he take  $\beta$  to imply (or at least help to imply) the principle he is trying to prove in the First Analogy as a whole?

### 3. The Possible Grounds of $\beta$

The sentence we have looked at from B225 presents the need for a substratum in the objects of perception that does not come to exist or cease to exist as a consequence of the fact that time cannot be perceived in itself. The original proof in the first edition begins with the assertion that our apprehension of the manifold given in or through sense is always successive, and thus always changing; having done so, it immediately goes on to draw an intermediate conclusion very similar to  $\beta$ .<sup>6</sup> The claim we have looked at from A188/B231 is followed by the affirmation that this very thing that persists is what makes possible the representation of the transition from non-existence into existence (or vice-versa). We can thus cite at least three considerations relevant to Kant's adoption of  $\beta$ . They can even be combined to form a syllogism:

- (1) Our apprehension of the manifold given in or through sense is always successive, and thus always changing.
- (2) Time cannot be perceived in itself.

---

<sup>6</sup> The relevant passage is cited in note 3, above.

(3) Therefore, what already existed and continues to exist, i. e., the substratum in the objects of perception, makes possible the representation of the transition from non-existence into existence (or vice-versa).

The truth of  $\beta$  would follow from that of (3), but it is not at all clear that or how the truth of (3) would follow from that of (1) and (2). We thus still lack an answer to our first question: Why would Kant subscribe to  $\beta$ ? He apparently tries to explain his reason in the incredibly opaque and confusing passage that follows the assertion of (3) on A188/B231, but the passage in question turns out (at least in my view) to be a dead end.<sup>7</sup> The best way to find an answer, I therefore believe, is to wrestle independently with the more basic question that Kant himself raises. Let us do so in terms of a house of the sort often built by children out of blocks.<sup>8</sup> How can we perceive that this house has come to exist? Under what conditions can we be aware, on the basis of apprehension, that it has done so?

The first such condition, it seems clear, is that we apprehend the house. Let us say that we do this at time<sub>2</sub>. What we apprehend at this time can be regarded in a number of ways. It can be regarded as a house, of course, but it can also (and equally well) be regarded as a set of blocks, a set of molecules, a set of atoms, or whatever.

The second such condition is that we remember having earlier (at time<sub>1</sub>) apprehended a state of affairs other than the one we apprehend at time<sub>2</sub>.

It should be clear at first glance why these conditions must be met. Equally clear upon reflection, however, is that the meeting of these conditions alone will not be sufficient. In many cases, after all, these conditions are both met but we are not aware that the house in question

---

7 “If you assume that something simply began to be, then you would have to have a point of time in which it did not exist. But what would you attach this to, if not to that which already exists. For an empty time that would precede is not an object of perception; *but if you connect this origination to things that existed antecedently and which endure until that which arises, then the latter would be only a determination of the former, as that which persists.* It is just the same with perishing: for this presupposes the empirical representation of a time at which there is no longer an appearance.” What remains completely unexplained in this passage from A188/B231 is the ground of the assertion I have placed in italics.

8 This example is adapted from D. P. Dryer, to whom much of the analysis that follows in this section is greatly indebted. See D. P. Dryer, *Kant's Solution for Verification in Metaphysics* (Toronto: University of Toronto Press, 1966), 353–9 in particular.

has come to exist. The meeting of these two conditions can result in that awareness only if we take what we remember apprehending at time<sub>1</sub> to be incompatible with the existence of the house at that time. But what could we remember apprehending at time<sub>1</sub> and take to be incompatible with the existence of the house at that time? The only possible answer, one might well suggest, is something that makes up the house at time<sub>2</sub>. If what we remember apprehending at time<sub>1</sub> is a set of blocks scattered all over the floor, for example, and we take that to be the same set that now makes up the house, we will take what we remember apprehending at time<sub>1</sub> to be incompatible with the existence of the house at that time. As a result of so doing, moreover, we will at least *seem* to be aware that the house has come to exist. At least *seeming* to be aware that this has happened is a necessary condition of *being* aware that it has happened, of course. What follows (on the above suggestion) is that we can be aware that the house has come to exist, on the basis of such apprehension, only if we take what makes it up to be what we remember apprehending in a different form at an earlier time. What we are implicitly doing when we do that, however, is taking the resulting coming to exist that we at least seem to be aware of to be a mere change in the determinations of something that already existed and continues to exist. The conclusion we eventually end up with, therefore, is that we can be aware, on the basis of apprehension, that the house has come to exist only if we take its doing so to have been a mere change in the determinations of something that already existed and continues to exist.

A similar analysis could easily be offered in regard to a hypothetical case where we are aware, on the basis of apprehension, that the house has ceased to exist. The key issue is whether a similar analysis both could and *must* be offered in *all* cases where we are aware, on the basis of apprehension, that something has come to exist or ceased to exist. This issue is large and cannot be explored here. My tentative conviction, however, is that the answer may be yes.<sup>9</sup> Just for the sake of ar-

---

9 I was convinced for a very long time that the answer is no. The reason had to do with the way this conclusion is reached by some of the commentators who rely on it in their own interpretations. Both Dryer and Longuenesse appear to think it is grounded on the more general claim that we can be aware of an event or case of objective (as opposed to merely subjective) succession only if we take it to be a mere change in the determinations of something that already existed and continues to exist. That more general claim seems untenable to me, so I took it for granted that the more specific claim that they seem to base on it



gument, therefore, let us assume that it is. The result is then a claim that might best be dubbed  $\alpha$ :

$\alpha$ : We can be aware, on the basis of apprehension, that something has come to exist or ceased to exist only if we take its doing so to have been a mere change in the determinations of something that already existed and continues to exist.

This claim is similar but not identical to  $\beta$ . It is a claim about what *we* must *do* if we are to be aware, on the basis of apprehension, that something has come to exist or ceased to exist;  $\beta$  is a claim about what must *in fact be the case* if we are to be aware, on the basis of apprehension, that such a thing has occurred. We have just seen why Kant might subscribe to  $\alpha$ . We can even imagine why he might regard it as too obvious to stand in need of explanation. Might a commitment to  $\alpha$  be what grounds his adoption of  $\beta$ ?

It could do so only in conjunction with some further claim or principle. Such a further claim or principle, however, might seem easy to supply: "... it is an analytic truth that any object represented must conform to the conditions under which alone it can be represented as an object," asserts Henry Allison.<sup>10</sup> Let us dub this AP (for "Allison's Principle"); at least for the sake of argument, furthermore, let us assume that it is true:

AP: Any object represented must conform to the conditions under which alone it can be represented as an object.

Our apprehension, insists Kant, is always successive and thus always changing. In our apprehension, that is to say, the elements of the manifold given in or through sense are *always* coming to exist and ceasing to exist. In some but not all cases, nonetheless, we at least seem to be aware that such an element really comes to exist or ceases to exist, that it does so not only in apprehension but also "in the object". What happens in such cases is that something (namely, a case of coming to exist or ceasing to exist) is represented as an object. But according to  $\alpha$ , it seems clear,

---

must be so as well. Henry Allison may avoid this conflation for the most part, but there are nonetheless passages in his chapter on the First Analogy that seem to imply it. The most striking is on page 205 of Henry Allison, *Kant's Transcendental Idealism: An Interpretation and Defense*, first edition (New Haven and London: Yale University Press, 1983), where what he describes as "replacement change" seems equivalent to an event or case of objective as opposed to merely subjective succession in general.

10 Allison, *Kant's Transcendental Idealism*, 29.

this can happen only if we take the case of coming to exist or ceasing to exist in question to be a mere change in the determinations of something that already existed and continues to exist. Our doing this, in short, is (at least according to  $\alpha$ ) a condition under which alone something (namely, a case of coming to exist or ceasing to exist) can be represented as an object. What this in turn entails, according to AP, is that any case of coming to exist or ceasing to exist that is represented as an object (i. e., as objective) must conform to that condition. But how could it do so unless it were in fact what we must take it to be in order to represent it in this way? Assuming that this is indeed the only way that it could do so, the implication of  $\alpha$  and AP is none other than  $\beta$ .

#### 4. The Path from $\beta$ to Kant's Conclusion

We now have an answer to the first of our questions: Why might Kant subscribe to  $\beta$ ? We still need an answer to the second, however: Why would he take  $\beta$  to imply (or at least help to imply) the principle he defends in the First Analogy as a whole?

The proper interpretation and ultimate tenability of Kant's views on space and time are of course controversial. In some passages, however, he undeniably seems to imply that all events in general, and thus all cases of coming to exist and ceasing to exist in particular, are cases we could (at least in principle) be aware of on the basis of apprehension. Some of the most striking such passages are found in Section Six of the antinomy of pure reason, a section entitled "Transcendental idealism as the key to solving the cosmological dialectic." Let us call this contention TI, since Kant apparently takes it to be implied by his transcendental idealism:

TI: All cases of coming to exist and ceasing to exist are cases we could (at least in principle) be aware of on the basis of apprehension.

The addition of TI leads straight to a claim that goes further than  $\beta$ . In the absence of TI, it would be hard to deny that cases of coming to exist or ceasing to exist might occur that we cannot and/or could not (even in principle) be aware of on the basis of apprehension. In the absence of TI, therefore, Kant's adherence to  $\beta$  would give him no reason to infer that *every* case of coming to exist or ceasing to exist must be a mere change in the determinations of something that already existed and continues to exist. With the addition of TI, on the other hand, his adher-

ence to  $\beta$  does put him in a position to infer that, and thus to insist on the claim we shall henceforth call  $\gamma$ :

$\gamma$ : Every case of coming to exist or ceasing to exist is a mere change in the determinations of something that already existed and continues to exist.

It may perhaps seem that  $\gamma$  is still a long way from the principle Kant aims to defend. In reality, however, it is not all that far. Even on its own,  $\gamma$  seems to imply that whatever comes to exist or ceases to exist is either (a) a determination, or (b) something whose existence consists in the exemplification of determinations by something else.<sup>11</sup> The concept of a substance, holds Kant, is that of something that could exist only as subject but never as predicate or determination (B149). This must surely be regarded as an oversimplification on his part, however. Lots of things come to exist and cease to exist but are not mere predicates or determinations. My fist is an example; so are tables, chairs, apples, and so on.<sup>12</sup> An essential part of what he is trying to prove, however, is that no substance ever comes to exist or ceases to exist. His considered position, therefore, must be that the distinction between a substance on the one hand and a determination on the other is not exhaustive. There must be a third slot, so to speak, for things that are not determinations but also come to exist and/or cease to exist. The obvious candidate is one filled with things whose existence consists in the exemplification of properties or determinations by something else. It makes no apparent sense to say that a desk or an apple is a property or determination, after all, but it makes perfectly good sense to say they are things whose existence consists in the exemplification of properties or determinations by something else (sets of molecules or atoms, for example).

Kant's considered position, it thus seems, must be that the concept of a substance is that of something that (a) is a subject, (b) is not a determination, and (c) is not something whose existence consists in the exemplification of determinations by something else. Given this account of that concept, however, there is a hopeless clash between  $\gamma$  on the one hand and the supposition that something that comes to exist or ceases to exist is a substance on the other. If there is substance at all,

---

11 How could the coming to exist or ceasing to exist of  $x$  be a mere change in the determinations of  $y$  unless  $x$  is either (a) a determination of  $y$ , or (b) something whose existence consists in the exemplification of certain determinations by  $y$ ?

12 The following account of what Kant must really mean by "substance" is based on that of James Van Cleve. See James Van Cleve, *Problems from Kant* (New York and Oxford: Oxford University Press, 1999), 105.

given  $\gamma$ , it can only be something that has never come to exist and will never cease to exist.

What cannot be established on the basis of  $\gamma$  alone is that there is substance at all and that whatever comes to exist or ceases to exist is either a determination of substance or something whose existence consists in the exemplification of certain determinations by substance. The most that can be shown on its basis *alone* is that there is something that *might be* substance and that whatever comes to exist or ceases to exist is either a determination of something that *might be* substance or something whose existence consists in the exemplification determinations by something that *might be* substance.<sup>13</sup> (The explanation of why this is so is quite complex and convoluted; it has therefore been relegated to note 13, below.) In order to establish the *former* claim, and thus to

---

13 In order to show this, let us assume that  $\gamma$  is true and that something (call it  $x$ ) comes to exist. The coming to exist of  $x$  is then a mere change in the determinations of something else (call it  $w$ ) that already existed and continues to exist in the form of  $x$ .  $X$  is either (a) a determination of  $w$ , or (b) something whose existence consists in the exemplification of certain determinations by  $w$ .  $X$  is therefore not a substance.  $W$  might be, for all we know. It has not ceased to exist and may not have come to exist either.  $X$ , therefore, is either (a) a determination of something that might be substance, or (b) something whose existence consists in the exemplification of determinations by something that might be substance. Let us now assume that  $w$  has come to exist after all. The coming to exist of  $w$  was then a mere change in the determinations of something else (call it  $v$ ) that already existed and continues to exist in the form of  $w$  (and is still around in the form of  $x$ ).  $W$  is either (a) a determination of  $v$ , or (b) something whose existence consists in the exemplification of certain determinations by  $v$ .  $W$  is therefore not a substance.  $V$  might be, for all we know. It has not ceased to exist and may not have come to exist either.  $W$ , therefore, is either (a) a determination of something that might be substance, or (b) something whose existence consists in the exemplification of determinations by something that might be substance. Since this is the case, moreover, and since  $x$  is either (a) a determination of  $w$ , or (b) something whose existence consists in the exemplification of certain determinations by  $w$ ,  $x$  too is still either (a) a determination of something that might be substance (namely  $v$ ), or (b) something whose existence consists in the exemplification of certain determinations by something that might be substance (namely  $v$ ). (If  $x$  is a determination of  $w$  and  $w$  is a determination of  $v$ , I am assuming,  $x$  is also a determination of  $v$ .) No matter how far we pursue this, it should be clear, we will be left with something that is still around in the form of  $x$ ,  $w$ , etc., and may not have come to exist. No matter how far we pursue this, in other words, we will be left with something that might be substance. Let us call this SC, for "substance candidate". All members of the chain further up will be either (a) determinations of SC, or (b) things whose existence consists in the exemplification of certain determinations by SC.

reach the principle Kant is aiming at, we need one further link. What the link in question must state is that there cannot be an infinite downward series of determinations and/or things whose existence consists in the exemplification of determinations by something else.<sup>14</sup> Let us call this link ARP (for “anti-regress premise”):

ARP: There cannot be an infinite downward series of determinations and/or things whose existence consists in the exemplification of determinations by something else. Any such series must be grounded at some point in something that truly qualifies as substance.

The addition of ARP to  $\gamma$  finally leads to the principle Kant is trying to prove in the First Analogy as a whole. Given the truth of both ARP and  $\gamma$ , there must be something that truly qualifies as substance and all entities of the sort that come to exist and/or cease to exist (i. e., “appearances”, in Kant’s language) must be either (a) determinations of it, or (b) things whose existence consists in the exemplification of certain determinations by it. The substance in question cannot be something that exists apart from the entities in question (albeit in the sphere of such entities); it must be something *within* them, something they *contain*. In all change of such entities (i. e., in all coming to exist and/or ceasing to exist of them), the substance within them must persist.

## 5. Assessment

In this short essay we cannot assess this proof in detail. The truly critical steps, it seems clear, are  $\alpha$ , AP, TI, and ARP; the truth of  $\beta$  would follow from that of  $\alpha$  and AP, the truth of  $\gamma$  would follow from that of  $\beta$  and TI, and the truth of the principle Kant aims at in the First Analogy as a whole would follow from that of  $\gamma$  and ARP. The fourth of these truly critical steps has great intuitive plausibility. Even if it could not be established on a more solid basis, moreover, the implications of  $\gamma$  by itself are so far-reaching that the difference between them and the principle Kant aims at might not seem very crucial. An assessment of the third would have to be prefaced by an account of Kant’s transcendental idealism in general and cannot even be broached here. The second

---

14 To the best of my knowledge, the only major commentator to have explicitly recognized and acknowledged the need for this link is Van Cleve. See Van Cleve, *Problems from Kant*, 109–11.

seems suspicious.<sup>15</sup> It is clearly analytic in some sense but I am not at all sure that it is analytic—or even true—in the sense that would be required to license the move from  $\alpha$  to  $\beta$ . The first, however, is of very great interest no matter what our eventual verdict on the others. It forms part of what I take to be Kant's overall solution to a set of problems no one before him had explored, or perhaps even seen. The problems in question concern the possibility of objective time determination: how we can be aware, or even *seem* to be aware, of a distinction between objective and merely subjective in regard to temporal relations and occurrences. Kant's recognition of these problems and proposed solutions to them, as presented in all three of the analogies, stand among his most original insights and greatest contributions to the attempt made in theoretical philosophy to explain how human persons obtain a unified knowledge of objects and even the very conception of a distinction between objectivity and subjectivity.

---

15 Basically, it seems to me that something might *conform* to a concept in the sense of allowing itself to be subsumed under it without actually *corresponding* to it. What is needed to license the move from  $\alpha$  to  $\beta$ , however, is a claim to the effect that any object represented must *correspond* to any concept whose application by us allows it to be represented as an object. This does not seem obvious to me.

## 5. Kant's Attack on Leibniz's and Locke's Amphibolies

*Andrew Brook*

### 1. Introduction

The Transcendental Analytic of Kant's *Critique of Pure Reason* ends with a little appendix called The Amphiboly of the Concepts of Reflection. It contains a number of important discussions.

First, the passage contains the only discussion of numerical identity in the entire first *Critique*. Kant had used the concept earlier a number of times; but he discusses it for the first (and only) time here.<sup>1</sup> In the Table of Categories, numerical identity is not even mentioned—an omission that is at least puzzling when we see Kant's discussion of the concept here.

Second, the passage offers not one but two arguments for the need for sensible intuition—the first such arguments anywhere in the Analytic. Kant has of course asserted this need earlier, many times, in fact—but try to find an argument! The arguments that he offers here were not new to him. Indeed, they go back to his earliest philosophical writings. One of the two builds on his discussion of identity.

Third, the passage introduces a new form of transcendental thinking, transcendental reflection.<sup>2</sup> Transcendental reflection, Kant tells us, is one route to synthetic *a priori* knowledge, so it should have been important to him. Yet it had never appeared in his work prior to this Ap-

---

1 The term “identity of number” is perhaps preferable to “numerical identity”—identity comes in only one kind—but I will use Kant's term (A263/B319).

2 There is an interesting terminological issue here. Kemp Smith translates both “*Reflexion*” and “*Überlegung*” as “reflection”, yet the words have different meanings. Kant uses the first in the phrase “Concepts of Reflection”, the second in the name for the new transcendental activity that Kemp Smith labels “Transcendental Reflection”. Pluhar suggests that “deliberation” is a better translation of “*Überlegung*” than “reflection”. He may have a point. However, since Kant parses “*Überlegung*” in Latin as “*reflexio*”, I will stick with Kemp Smith's now standard translation.

pendix, not under this name anyway—the importance of this qualification will become clear later—and immediately disappeared again.<sup>3</sup> Altogether, a curious little piece of work.

The Appendix has not, to say the least, fired the imagination of Kant's readers. With the possible exception of the Transcendental Doctrine of Method, no part of the first *Critique* has received less attention. This neglect is not warranted.<sup>4</sup>

## 2. The Attack on Locke and Leibniz

Kant presents the Appendix as an attack on two mistakes, one made by Leibniz and one by Locke, mistakes that are in some ways the opposite of one another. An amphiboly of the concepts of reflection, Kant tells us, is “a confounding of an object of pure understanding with appearance” (A270/B326).<sup>5</sup> This confounding can happen in two ways. One is to take an appearance to be an object solely of the understanding. Here one takes something that has in fact been delivered through the senses (or sensible imagination, presumably) to be something known purely by thinking about it. This is the form Leibniz's amphiboly took. The other is to take something acquired nonsensibly (e. g., *a priori* concepts such as the categories) to be a deliverance of the senses. This is the route Locke took. “In a word, Leibniz *intellectualised* appearances, just as Locke ... *sensualised* all concepts of the understanding, i. e., interpreted them as nothing more than empirical or abstracted concepts” (A271/B327).

---

3 Almost disappears. Kant mentions the idea again once in the opening paragraphs of the Dialectic immediately following and once more in the *Prolegomena* (4:326).

4 Norman Kemp Smith, *Commentary to Kant's "Critique of Pure Reason"* (London: Macmillan & Co. Ltd, 1923); G. H. R. Parkinson, “Kant as a Critic of Leibniz: The Amphiboly of Concepts of Reflection”, *Revue Internationale de Philosophie* 35 (1981); R. Butts (ed.), *Kant's Philosophy of Physical Science* (Boston: Reidel, 1986); Derk Pereboom, “Kant's Amphiboly”, *Archiv für Geschichte de Philosophie* 73 (1991), 50–70; K. Ameriks, “The Critique of Metaphysics: Kant and Traditional Ontology”, in P. Guyer (ed.), *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press, 1992), 249–79; and J. Wuerth, “The First Paralogism, its Origin, and its Evolution: Kant on How the Soul Both Is and Is Not a Substance” (ch. 6 in the present volume).

5 Quotes from Kant's first *Critique* are from *Immanuel Kant's Critique of Pure Reason*, tr. Norman Kemp Smith (London: Macmillan Co. Ltd., 1963 [1927]).



Were Leibniz right, Kant thought, all genuine knowledge would be purely conceptual, a product of acts of the understanding; sensible experience would have no essential role to play in determining the truth or falsity of beliefs. For Kant, this would be wrong in itself. Even worse, however, if all knowledge were conceptual, Kant believed, we would or at least could have knowledge of things as they are in themselves. Why? Because with objects of understanding, their nature (as they are, not just as they appear) would be carried by their concepts. What concepts carry is accessible to us, so if we know what we know via concepts, we would know these objects as they are. Kant, of course, had to deny that we have any such knowledge.

Locke's amphiboly makes the opposite mistake. As Kant read him, Locke held to the old Aristotelian idea that everything we know comes to us via the senses. This would be a serious mistake because it would undermine the claims of propositions in, for example, mathematics and physics to be necessary and universal, radically psychologizing them. Kant took such necessity and universality to be self-evident in the case of mathematics and required of anything with a claim to be a science (B20–1), so he could not accept Locke's empiricism any more than he could accept Leibniz's rationalism.

For Kant, any amphiboly is serious and Locke's psychologism would have struck him as utterly mistaken. If we are to judge by what follows, however, Leibniz's amphiboly concerned him a good deal more than Locke's. Locke is not mentioned again.

### 3. Transcendental Reflection and Its Concepts

How does Kant set out to refute these two amphibolies? He turns to what he calls transcendental reflection, "the consciousness of the relationship of representations to our different sources of knowledge" (A260/B316).<sup>6</sup> It uses what Kant calls the concepts of reflection to achieve this. The different sources of knowledge that he has in mind are sensibility and understanding, and the task of transcendental reflection is to determine the source of a given representation.

Kant's arguments against Leibniz and Locke fall out of this investigation. Leibniz mistook objects belonging to sensibility for objects re-

---

6 He actually refers to reflection without qualification, but the context makes it clear that he means transcendental reflection.

quiring only understanding (i. e., thinking) to be known. As Kant put it, Leibniz wrongly supposes that if a distinction is not found in the concept of a thing, then it is not to be found in the thing (A281/B337):

Because in the mere concept of a thing in general we abstract from the many necessary conditions of its intuition, the conditions from which we have abstracted are, with strange presumption, treated as not being there at all, and nothing is allowed to the thing beyond what is contained in its concept.

On the other hand, Locke took objects requiring acts of understanding to belong solely to sensibility.

The way to determine from which source of knowledge a represented object comes is to study the kinds of relations it enters into with other represented objects. Kant says that four kinds of relationship are germane: relations of identity<sup>7</sup> and difference, agreement and opposition, inner and outer, and matter and form. These are the concepts of reflection. Relations of identity and difference concern the conditions of multiple representations representing one and the same object—numerical identity of object across representations. Relations of agreement and opposition are about the very different ways different kinds of object can be in opposition to other objects. The distinction between inner and outer concerns some complex issues in Leibnizian metaphysics. The general issue behind the distinction is a particular form of the distinction between intrinsic and extrinsic properties. Finally, by the terms matter and form Kant had something broader in mind than what would occur to us now, something very much like the Aristotelian distinction between utterly formless stuff and the forms it takes when made into objects.<sup>8</sup>

Of the four pairs, identity and difference and agreement and opposition are the most important for our purposes here. The issues of enduring interest for Kant's project mentioned at the beginning all arise in the course of his treatment of these two pairs.

---

7 Kemp Smith translates Kant's word "*Einerleiheit*" as "identity", Pluhar as "sameness". Strictly speaking, Pluhar probably made the better choice. Since, however, Kant clearly has numerical identity in mind—four lines down he actually speaks of *numerica identitas* (A263/B319)—I will follow Kemp Smith.

8 Kant makes this matter/form distinction at A86/B118, A166/B207, and A267/B323.

#### 4. Identity and Difference

In connection with identity and difference, the question that interests Kant most is the relation of identity of qualities (i. e., indiscernibility) to numerical identity (i. e., being one and the same thing; *numerica identitas*, as he puts it) (A263/B319). His main target is the principle of the identity of indiscernibles.

Contrary to Leibniz's principle, Kant urges, the objects of two representations can be entirely indiscernible, that is to say, qualitatively indistinguishable, yet be different objects (A264/B320): "the mere fact that they [the two represented objects] have been intuited simultaneously in different spatial positions is sufficient justification for holding them to be numerically different." He gives the now famous example of drops of rain: two drops of rain are clearly *two* drops. Yet they can be indiscernible. Under what condition? When one description is all we need to describe both fully. Since what makes a drop of water the thing it is does not change when it changes location, location is not part of what it is (cf. A272/B328). Now, if the two drops are descriptively indiscernible, the only way we could be aware of their numerical difference, urges Kant, is by sensibly intuiting them. QED.

As Kant saw it, indiscernibility entails identity only with respect to objects of *the understanding*; Leibniz's mistake was to hold that the same is true of objects of the senses. To misapply the principle in this way is to commit an amphiboly, confusing the relations distinctive to objects of the understanding with the relations distinctive to objects known via the senses.

#### 5. Agreement and Opposition

The distinction between understanding (thinking) and sensibility is central to Kant's analysis. What difference, exactly, is marked by this distinction? The next pair of concepts, agreement and opposition, helps clarify this difference. Relations of agreement and opposition are about the ways in which represented objects can be in opposition. The relevant relationships fall into two broad kinds: semantic and non-semantic.

The relations of opposition between objects such as numbers and propositions are semantic: inconsistent implication, semantic exclusion, disjunction, contradiction, etc. How Kant saw such relationships is none

too clear, being enmeshed in Leibniz's principle that, as Kant put it, "realities ... never logically conflict" (A273/B329). However, the contrast Kant wanted to make is clear, so this may not matter much.

As well as semantic properties, many objects also have (or are represented as having) temporal properties and some have both spatial and temporal ones. This opens the door to additional, nonsemantic forms of opposition. Consider, for example, the way one force can counteract another and the way feelings of pain "counterbalance" feelings of pleasure. In Kant's words (A265/B321),

When such realities are combined in the same subject, one may destroy the consequences of another, as in the case of two moving forces in the same straight line, in so far as they either attract or impel a point in opposite directions, or again in the case of a pleasure counterbalancing pain.

The relationships resulting from spatial properties such as mass, energy, size, color, texture, and so on, that open the way to these forms of opposition are nonsemantic.

Kant's anti-Leibnizian conclusion follows immediately. We could be aware of nonsemantic forms of opposition only if we have sensible awareness of objects. We are aware of these forms of opposition. Therefore, ... QED. While we can know objects of the understanding purely by their semantic properties, so by merely thinking about them, we need sensibility to know objects that enter into nonsemantic relationships.

Note two things. First, Kant has not proved his claim about the need for "sensible intuition" generally, only for the cases he discusses. Second, his earlier point about identity could now be put this way: for some objects, nonsemantic properties play a crucial role in numerical identity and difference.

The argument for the necessity of sensible intuition implies in turn, Kant thought, that we are not aware of anything for which we require sensible intuition as it is in itself. Sensible intuition tells us merely how things appear to us, not how they are. For any object for which we require sensible intuition, then, we are not aware of that object as it is in itself. Once more, QED.

The issues raised by the remaining two pairs of concepts of reflection, inner/outer and form/matter, are less important to Kant's overall

project than those raised by the first two pairs. In this short essay, I will say no more about them.<sup>9</sup>

## 6. Numerical Identity and the Table of Categories

Kant's arguments in connection with the first two pairs of concepts of reflection relate to the rest of the Analytic rather curiously. Let us consider identity and difference first.

Remarkably, Kant's observations about numerical identity in the Appendix constitute the first and only discussion of the concept in the whole first *Critique!* The concept is not part of the Table of Categories and Kant does not discuss it anywhere else. "But he used the concept of identity over and over!", it will be objected. He did indeed; how could he not? It underlies the synthesis of recognition. In the A-edition he even uses the word "identity" in this connection once (A115). He refers explicitly to the identity of consciousness a number of times (e.g., A108, A113). He speaks of the notion of a unit, a quantity of one, in connection with the synthesis of apprehension (e.g., B162) and of the "successive apprehension of an object" (A145/B184) in the Schematism. He talks of the "identity of the substratum" through change in the First Analogy (A186/B229): for change to be possible, some unit of something must persist through the change. And in the Second Analogy he lays out some of the conditions of apprehending an object as a single persisting object (A198/B243). In short, prior to the Appendix, Kant makes extensive *use* of the concept of numerical identity. He makes even more direct use of it and by name after the Appendix in the critique of the second and third paralogisms (e.g., A362). Despite using the concept of identity so frequently, however, Kant nowhere *discusses* it until the Appendix. This is more than a little strange. What could be more basic to representation of objects or a theory of synthetic knowledge generally than individuation and re-identification? Surely they are at least as basic as, say, modality. Strange to be sure; but true.

Concerning the Table of Categories, the first question that comes to many people is, "Why is numerical identity not part of it?" The ques-

---

9 I hope to say more about the final two pairs of concepts of reflection in a longer version of this paper. Derk Pereboom's excellent paper, "Kant's Amphiboly", is entirely devoted to the inner/outer pair.

tion must have worried Kant because he suddenly takes it up for no discernible reason in the Appendix. There he tells us that identity and difference and the other concepts of reflection “are distinguished from the categories by the fact that they do not present the object according to what constitutes its concept (quality, reality), but only serve to describe in all its manifoldness the comparison of the representations which is prior to the concepts of things” (A269/B325). The same issue again pops up out of nowhere in the *Prolegomena* (the only place concepts of reflection are ever mentioned after the first *Critique*). There Kant repeats that we must not confuse concepts of reflection and categories; categories apply to objects, “whereas the former are only concepts of a mere comparison of concepts already given and therefore are of quite another nature and use” (4:326).

The two passages present a problem. In the first passage, Kant seems to suggest that we apply the concept of identity and the other concepts of reflection *prior* to application of the categories and representation of objects, whereas in the second, the concepts of reflection seem to come in *after* application of the categories. Yet neither suggestion is consistent with the way he himself used the concept of numerical identity in the Synthesis of Recognition in a Concept (found in both editions). There he insisted that not just retaining but also re-identifying an object as one presented earlier (i. e., recognizing that an object “is the same as what we cognized a moment before”) is *part of* recognizing it as an object. Yet we recognize it as an object via the application of categories (A103; see B130–1). If the Table of Categories has to include number, qualities, relations, and modality (existence status), surely it should also include a thing being one thing and remaining the same thing over time.

In the precritical writings, Kant's name for what he called analytic relationships in *CPR* was the “law of identity” (we will return to it). It would have been part of General Logic. Since the Categories are part of Transcendental Logic, if Kant thought that numerical identity is in General Logic, that would have been a good reason for leaving it off the Table of Categories (and relegating the concept to an appendix).<sup>10</sup>

If this were Kant's view (and there is not a lot of evidence one way or the other), he would have been wrong by his own lights. Kant's clearest specification of transcendental logic is that “it concerns itself with understanding and reason solely insofar as they relate *a priori* to ob-

---

10 Stephen Palmquist raised this possibility.

jects” (A57/B81). Applying the concept of numerical identity to experience is applying a concept to experience, one moreover that would appear to be as necessary for experiencing objects and therefore as *a priori* as any of the ones that did make the Table.

Another reason for leaving numerical identity off the Table may have been this. Kant saw logic, specifically the Aristotelian forms of judgment, as the template for the categories. To generate the categories, he just filled out the forms of judgment with some content. That this move forces the exclusion of numerical identity shows, I think, how bad a move it was.

In connection with the Table of Categories, we have focused on the concept of identity but a similar question could be asked about the other concepts of reflection: agreement and opposition, inner and outer, and matter and form. Is it not just as necessary and just as transcendental (just as much a matter of relating *a priori* concepts to objects of experience) that an object of representation be compatible with some things and in opposition to others as that it be individuated and re-identifiable? Or that it have inner and, in most cases, outer properties? Or that it consists of some kind of informed material? Well, maybe; but at least opposition, the outer, and form, being relational concepts, can be accommodated within the categories that Kant allowed under the heading of Relation.

## 7. Concepts of Reflection and Sensible Intuition

Kant’s remarks in the Appendix about identity are interesting in their own right. In addition, as we saw, they form the basis of one of his two arguments there for the need for sensible intuition. The other arises from his treatment of agreement and opposition. Kant had argued for this need, as is well-known, in the *Inaugural Dissertation* of 1770 and he reaffirmed it repeatedly in *CPR*. But these are the first arguments for the claim anywhere in *CPR*. They come very late. Kant argues for the need for the *forms* of sensible intuition, space and time, in the *Aesthetic* and he identifies sensibility as one of the “two stems of human knowledge” as early as the Introduction (A15/B29) and the first paragraph of the *Aesthetic* (A19/B33). Indeed, in the Preface to the B-edition (Bxxv–vi), he says explicitly that intuition is *necessary* for knowledge and tells us he will prove the assertion in “the analytic part of the *Critique*”. He urges that we need sensible intuition again a

number of times in the first two paragraphs of the Transcendental Logic, saying, for example, “without sensibility no objects would be given to us” and, famously, “thoughts without content are empty” (A51/B74), and repeatedly thereafter throughout the *Analytic*. The trouble is, not once does he offer an argument for it—until he gets to the Appendix.<sup>11</sup>

This might not be as surprising as it appears. Ten years elapsed between the *Inaugural Dissertation* and the final assembling of *CPR*. That was more than enough time for Kant's interests to have shifted. Whereas the role of sensible intuition was front and centre in the *Dissertation*, in *CPR* Kant's focus is first of all on the shape, justification, and limits of the nonsensible, the *a priori* component of knowledge. The Introduction, for example, is entirely devoted to the sources of *a priori* and yet synthetic knowledge. When sensible intuition does appear, as it does at the very beginning and the very end of the Introduction and in the first paragraphs of the *Aesthetic*, Kant advances no argument for his claim that we need it. He acts like he is entitled to lay down the idea without argument. That was not how he proceeded in the *Inaugural Dissertation*.

By contrast, the *Dialectic* can be viewed as a single extended argument for the claim that the nonexperiential component cannot yield knowledge without experience. But it comes even later in the book.

In short, the absence in the *Analytic* of any argument for the claim that sensible intuition is required for knowledge is striking. The place where the need for such an argument is most pressing is in the opening pages of the Transcendental Logic, where Kant makes his famous claim, “thoughts without contents are empty” (A51/B75). What are Kant's reasons for saying this? He offers none. From then on, he acts as though the claim has been established and simply plugs it into other arguments and analyses as needed.

Two obscure passages in the sections immediately prior to the Appendix, the General Note to the Principles (occurring only in the B-edition) (B288) and the chapter on Phenomena and Noumena (A240/

---

11 The lack of an argument for the need for sensible intuition may connect to another gap. Kant never tells us what controls the contents of sensible intuition—needed if some experiences are to adjudicate beliefs (e.g., perceptions) and others not (e.g., dreams). I discuss this latter gap in “Critical Notice of L. Falkenstein, *Kant's Intuitionism: A Commentary on the Transcendental Aesthetic*”, *Canadian Journal of Philosophy* 29 (1998), 247–68.



B299), appear to contradict what I just said. However, (a) like the Appendix, they too come right at the end of the *Analytic* and (b) they merely anticipate the analysis in the Appendix and cannot be understood without it.

By contrast, the Appendix on the Amphiboly, as we have seen, does offer not one but two arguments for the claim. Most commentators focus on the drops of water argument. At first glance, it appears to be decisive. By itself, however, it is not. What it shows is that spatial location plays a role in some judgments of identity. But spatial location is part of one of the two *forms* of intuition. The example of the drops of water is no argument by itself that we need sensible *contents*.<sup>12</sup> However, if we add some moves that Kant would have found plausible, it can be turned into such an argument. We just need to add that there can be no space without spaces and no spaces without spatial content.

The argument from agreement and opposition works as it stands. As we saw, Kant argues that we could be aware of the way pain cancels pleasure, the way one force can counteract another, and so on, only via sensible experience. This is a powerful argument. That these arguments appear so late in *CPR* and in an appendix at that is puzzling.

## 8. History of These Arguments in Kant

Their late appearance in *CPR* was not because either argument was new to Kant. In fact, both arguments go all the way back to his first purely philosophical work, *Dilucidatio* (1755). Given that Kant offers no argument for the assertion that sensible experience is needed when he advances it early in *CPR*, it is puzzling that he did not at least bring these old arguments back into service—a perplexity only strengthened when we notice that he does offer an argument, indeed a related argument, for the claim at the appropriate place in the *Prolegomena*, the argument from incongruent counterparts (e.g., left- and right-handed gloves) (4:285–6). Let us look at the history.

Here is how the argument from indiscernible non-identicals appears in *Dilucidatio*. Says Kant, “things which are distinguished ... in virtue of space are not one and the same thing” (1:409). That is, indiscernibles having different spatial locations are not identical. As for nonsemantic opposition, Kant discusses forces and collisions at some length and clear-

---

<sup>12</sup> I owe this observation to Lorne Falkenstein.

ly articulates the contrast between what we have called nonsemantic and semantic oppositions (1:407–8). To be sure, Kant seems not yet to have noticed the potential of these arguments to create problems for Leibniz. Indeed, he treats them as compatible with Leibniz's epistemology. But they are there.

By 1763, Kant was explicitly directing similar arguments against Leibniz. As Kant read him, for Leibniz all real knowledge consists in uncovering analytic truths (i. e., spelling out in predicates what is contained in the subject of those predicates). To do this spelling out, we use what Kant calls, as I said, *the law of identity* (he means the content of the predicate being identical to at least part of the content of the subject, not numerical identity of objects). In the 1763 work, *Negative Magnitudes*, he goes after this nonempirical, semantic-relation-based theory of knowledge. With objects in space, he urges, “the motive forces of one and the same body which tend in exactly the opposite directions [can be] opposed to one another; [they] cancel their reciprocal consequences, namely, the motions” (2:193). Similarly for colliding objects, amounts of money, ships on multi-leg trips, and others (2:171–8). Such “real oppositions” are utterly different from the “logical oppositions” of contradictory concepts or propositions (2:172; see also 2:194, where he even uses the example of pain cancelling pleasure). By the end of the work, Kant takes it as established that logical form and semantic relations could not be all there is to knowledge (2:202):

I understand very well how a consequence can be posited by a ground according to the rule of identity, because it is found contained in it by dissection of the concept. But how something follows from something else, yet not according to the rule of identity—that is something which I would be glad to be able to make plain.

He is clearly saying that awareness of real oppositions requires a non-conceptual activity of the mind. By 1770 he calls it sensibility (2:392 f).

The argument from nonidentical indiscernibles did not reappear in the works of the 1760s. However, a closely related argument does: the argument from incongruent counterparts (again, left- and right-handed gloves are an example). It is closely related because it too is an argument that a difference between two things of which we are clearly aware cannot be expressed in a description. Kant used this argument in both 1768 (in “The Ultimate Foundation of the Differentiation of Directions in Space” [2:383]), where it is not aimed at Leibniz, and 1770 (in the *In-*

*augural Dissertation* [2:403]), where it is. Though it does not appear in *CPR*, it did appear again later, in the *Prolegomena* (4:285–6).

So by 1770, indeed by 1763, Kant was firmly settled in his view that sensible intuition is required for at least some kinds of knowledge. He seems not to have reached his own mature position, that *both* sensible intuition *and* use of concepts are needed for most kinds of knowledge, until *CPR*; but his critique of Leibniz on the purely conceptual nature of knowledge was largely complete by 1770. In short, the absence of arguments for the need for sensible intuition in *CPR* up to the Appendix, and the relegating of the arguments just discussed (when they do appear) to an appendix, are peculiar. When Kant first claimed early in *CPR* that sensible intuition is needed for knowledge, why did he not at least re-introduce his earlier, pre-critical arguments?

## 9. What happened to Transcendental Reflection?

I will close with another puzzling question: Why does transcendental reflection<sup>13</sup> not appear anywhere earlier than our Appendix, neither in the first *Critique* nor pre-critically, nor ever appear again?<sup>14</sup> Transcendental reflection should be a promising Critical method.

The first move in the Critical project that Kant did pursue, the Critical project of the Analytic, is to ask (A94/B126, A95–6): What are the necessary conditions of experience? The question at the heart of transcendental reflection is precisely analogous: What are the sources of knowledge necessary for the objects of our representations to have the relationships to other represented objects that they have? So why did Kant make so little of the term?

Part of the explanation is that Kant did do the work in other places, both pre-critically and in the first *Critique*, but under other names. One notable example is the method for metaphysical inquiry sketched in the *Prize Essay* (1764) and detailed in the *Inaugural Dissertation* (1770). In particular, in the latter work Kant introduces what he calls the fallacy of subreption and invokes his new doctrine that space and time are the subjective conditions of sensibility to expose a key example of it. The fallacy of subreption is the mistake of predicating spatial and tem-

---

13 Or rather, transcendental “*Überlegung*”, the German word that I translate as “reflection”.

14 We noted two insignificant exceptions to this statement in footnote 3.

poral properties of intellectual objects (2:411–2). The method used to expose it is transcendental reflection by another name and the mistake thus exposed is an amphiboly, though a bit different from the one attacked in the Appendix. There, chez Kant, Leibniz denied that sensibility provides knowledge. Here Leibniz gives objects of the understanding spatio-temporal properties.

Similarly, in all three chapters of the Dialectic, Kant paid close attention to the sources of knowledge: knowledge of the self in the Paralogisms, of the world in the Antinomies, and of “all reality” in the Ideal. Even the term “subreption” makes brief appearances in all three chapters. In the first-edition Paralogisms it appears at A389 and A402, in the Antinomies at A509/B537, and in the Ideal at A582/3/B610/11 and A619/B647.<sup>15</sup> Kant says not a word by way of introducing the term in any of these places and seems to assume that we will know what he means by it. However, Kant never used the words “transcendental reflection” in any of these passages nor anywhere else.

In a *Reflexion*, R5552 (18:218), probably written in 1778/79 as *CPR* was being finished, Kant makes a revealing connection. He says that “Concepts of Reflexion (their Amphiboly)” “can lead to paralogisms”, then lists the four pairs of concepts of reflection. Though he crossed out “can lead to paralogisms”, he gives his standard definition of “paralogism” between the opening remark about concepts of reflection and the list of the concepts of reflection, so he clearly intended the link to stand. After he lists the concepts of reflection, he then goes on to discuss issues that became the content of the Antinomies and the Ideal. These links between the concepts of reflection and the subject-matter of the three chapters of the Dialectic raise an interesting possibility. By the time Kant had finished the Dialectic, perhaps he thought that the important work of transcendental reflection had been done, not under that name, of course. If so, he may have further thought that he could safely relegate any remaining, purely anti-Leibnizian points to an appendix.

If those were his thoughts, he had overlooked a better alternative. He could have: (1) moved the arguments for the need for sensible intuition to early in *CPR* where they belong; (2) introduced subreption properly and combined it with his account of transcendental reflection; and (3) moved the combined discussion to the opening pages of the Dialectic—where indeed he does mention transcendental reflection.

---

15 Later, the term appears two more times: A643/B671 and A792/B820.

I will close with a brief summary of the whole essay. I have examined five issues with respect to Kant's Appendix on the Amphiboly of the Concepts of Reflection: (1) Kant's treatment of identity and what he calls real opposition; (2) the relationship between the concept of numerical identity and the categories; (3) Kant's arguments for the necessity of sensible contents in knowledge (the first such arguments in the first *Critique*); (4) the novel strategy Kant used to generate these analyses, a strategy he calls transcendental reflection; and (5) the puzzling history of these themes in the pre-critical and early Critical philosophy.<sup>16</sup>

---

16 Particular thanks to Jennifer McRobert for historical leads. Thanks to audiences at the APA Eastern Division, University of Western Ontario, University of Maryland, University of Waterloo, and Carleton University.

PART II

Personhood as a Problem  
for Rational Metaphysics



## 6. The First Paralogism, its Origin, and its Evolution: Kant on How the Soul Both Is and Is Not a Substance

*Julian Wuerth*

### 1. The Structure of the First Paralogism

Kant's rejection of the rational psychologists' conclusion that the soul is a substance is well known, presented by Kant in the first *Critique's* chapter on the Paralogisms of Pure Reason, specifically its First Paralogism. A "logical paralogism", as Kant defines it in the Paralogisms, "consists in the falsity of a syllogism due to its form, whatever its content may be" (A341/B399). The chapter on the Paralogisms, however, concerns itself with a *transcendental* paralogism, where there is a "*transcendental* ground for inferring falsely due to its form" (A341/B399). That there is a "transcendental ground" for the false inference is key to understanding this paralogism, and we will return to what this transcendental ground is, later. But for now, suffice it to say that this transcendental ground is such that the resulting fallacy is, as Kant says, "unavoidable", even if we can, in the end, expose it, so that it is not "insoluble" (A341/B399). Kant accordingly tells us in the *Jäsche Logic* that this fallacy is not a "*sophism*", where "one intentionally seeks to deceive others through it"; rather, as a "*paralogism*", "one deceives *oneself* through it" (9:134–5).

What, then, is the syllogism of the First Paralogism and where is the fallacy? The syllogism, as stated formally in the second edition of the First Paralogism (with added formal abbreviations, to be explained below), is this (B410–1):

[M] What cannot be thought otherwise than as subject [P] does not exist otherwise than as subject, and is therefore substance.

[S] Now a thinking thing, considered merely as such [M] cannot be thought otherwise than as subject.

Therefore [S] it [P] also exists only as such a thing, i. e., substance.

As Kant explains elsewhere (A323/B379), the three chapters of the *Dialectic*—the Paralogisms, the Antinomies, and the Ideal—each proceed



according to syllogisms. Not only does Kant clump these under the heading of the Dialectic rather than—as for Baumgarten—that of “special metaphysics”, but he also rearranges the usual ordering of rational cosmology, rational psychology, and rational theology to reflect the order of the relational categories. Rational psychology is now treated first, because its syllogisms are said to be of the *categorical* sort, beginning with a major premise that provides a universal rule of the categorical sort. Here there is a “mediating mark” or “middle concept” (M) attached to a “predicate” (P), and this “mediating mark” provides what Kant terms the “condition” of the rule of the major premise. The minor premise also includes a “mediating mark” and subsumes a cognition of a subject (S) under this mediating mark (M), or condition, of the rule. The conclusion follows, attaching to the minor premise’s subject (S) the major premise’s predicate (P). As Kant summarizes it in his lectures on logic, the form of the categorical syllogism is accordingly MP, SM, therefore SP.

Kant’s Paralogisms, including the first, have been much discussed, and one common question is where the logical fallacies are supposed to be found. In the case of the first paralogism, for example, the syllogistic form MP, SM, therefore SP, is valid. In addition to Kant’s claim that there is a transcendental ground for the fallacy, not a merely logical one, we have this clue to discovering the nature of the fallacy: Kant tells us that it lies in a *sophisma figurae dictionis* (A402) and that this sophism involves an equivocation between the transcendental and empirical uses of a term (A 402). Next, the mediating mark is the focus of Kant’s scrutiny, as evidenced by Kant’s remarks in his lectures on logic that, in a *sophisma figurae dictionis*, “the *medius terminus* is taken in different meanings” (9:135). The mediating term in the first paralogism is that of an absolute subject, or something that “cannot be thought otherwise than as subject” (B410). Accordingly, we would expect Kant to be arguing that the concept of an absolute subject is being used in two different senses, one transcendental and one empirical. On this reading, while this syllogism is valid at one level, because the mediating terms are identical *in name*, insight into the transcendental *meaning* of these identical terms reveals their difference and in turn the invalidity of the syllogism at this level. As Kant claims in the first edition, this syllogism’s conclusion is of the soul’s substantiality as understood in an empirical sense. This leaves us with the question of exactly what this conclusion of the soul’s substantiality is when understood in an empirical sense,

and how the *sophisma figurae dictionis* we employ to arrive at it is one that trades on different meanings whose difference is transcendental.

One further clue to bear in mind while addressing this question is the following, regarding the central role played by the concepts of immortality and permanence: Kant tells us in his lectures on metaphysics from the late 1770s, *Metaphysik L1*, that

The greatest yearning of a human being is not to know the actions of the soul, which one cognizes through experience, but rather its future state. The individual propositions of rational psychology are not so important here as *the general consideration of its origin, of its future state, and of survival* (28:263).

How does this goal of establishing our *immortality* relate to the conclusion of our *substantiality*? Kant's later *Metaphysik Dohna* (1792–3) makes the point that Kant often makes, that “For the future life is required: (1) the perdurability of the soul as substance, [etc.]” (28:688). Thus the whole point of rational psychology, in Kant's view, is to establish our immortality, and the conclusion of our substantiality is essential toward this end, meant to establish our *permanence*. Thus, when Kant asks in the First Paralogism, “But now what sort of use am I to make of this concept of a substance?” (A349), he answers that it has only one possible use, to establish our permanence, and that it fails in this:

That I, as a thinking being, *endure* for myself, that naturally I neither arise nor perish—this I can by no means infer, and yet it is for that *alone* that the concept of the substantiality of my thinking substance can be *useful* to me; without that I could very well dispense with it altogether (A349, emphasis added).

In sum, then, the First Paralogism includes a *sophisma figurae dictionis* of a transcendental sort concerning the concept of an absolute subject, leading the rationalists to their sought after empirical concept of substance in application to the soul as thing in itself, and the sole objective of this argument is to establish the permanence of this soul as thing in itself.

## 2. Kant's Conclusions on the Substantiality of the Soul as Thing in Itself

To highlight the target of Kant's attack in the First Paralogism, it will be helpful to review Kant's own conclusions that the soul is substance. The claim that Kant *does* conclude that the soul is substance would strike

many commentators as radical. This is the result in large part of understanding Kant's negative account of the soul against the rationalists in the Paralogisms as Kant's *complete* account of the soul. If we turn to Kant's broader recorded thought, however, we find that Kant regularly concludes the substantiality of the soul. By understanding exactly in what sense Kant embraces this conclusion, we can understand exactly in what sense Kant rejects this conclusion *as presented by the rationalists*. The pattern that emerges in the following quotes is this: Kant believes the soul as thing in itself is substance; that this concept of substance is that of a bare *substantiale*, namely of a *thing in general* that has powers enabling it to ground accidents; that because this thing in general is distinct from its accidents, it is devoid of all predicates; and, because it is distinct from all accidents, this *substantiale* is completely *indeterminate*. Finally, we see that for Kant this indeterminacy means we cannot assume this substance is permanent.

Without attempting to analyze how Kant reaches these conclusions, I will here present some of the relevant passages where Kant clearly affirms them:

(1) From the *Metaphysical Foundation of Natural Science* (1786):

And hence the clarity of the representations of my soul has such a degree, and in consequence of this fact the faculty of consciousness, namely, apperception—and along with this faculty even the *substance* of the soul—has also such a degree. But inasmuch as a total disappearance of this faculty of apperception must finally ensue upon the gradual diminution of the same, even the *substance* of the soul would be subjected to a gradual perishing, even though the soul were of a simple nature, because this disappearance of its fundamental force could not ensue through division (separation of substance from a composite) but, as it were, by expiration, and even this not in a moment, but by gradual remission of its degree, from whatever cause (4:542–3).

(2) From the first *Critique* (1781/1787):

Among the different kinds of unity according to concepts of the understanding belongs the causality of a *substance*, which is called “power.” At first glance the various appearances of one and the same *substance* show such diversity that one must assume almost as many powers as there are effects, *as in the human mind* there are sensation, consciousness, imagination, memory, wit, the power to distinguish, pleasure, desire, etc. (A648/B677–A650/B678, emphasis added).

(3) From *Reflexion* 6000 (1780–1789):

Whether the soul is a special substance? (18:420)

(4) From the well-known *Reflexion* 6001 (1780–1789):

The soul in transcendental apperception is *noumenal substance*; therefore no permanence of the same in time; and this can hold only for objects in space. (18:420–1)

(5) From *Metaphysik v. Schön* (late 1780s), after rejecting Spinoza's definition of substance and before rejecting Baumgarten's and Wolff's definition of substance:

Thus if I ask if something is a substance, I do not ask for a cause, but rather if it is in itself a thing.—If we in thought separate all accidents from a substance; in whom do these accidents inhere? The subject, which is their ground and that is called *substantiale*. The *substantiale* is an important concept of reason, but what is it actually? One often demands to know of the *soul*, what it actually is, not content with the accidents *inhering in it*, one demands also to know the subject of the same. In this the demand is unjust; for if I were to name this thus would I need to determine it according to accidents and precisely these are taken away. The *substantiale* is in general the concept of something, in which the accidents inhere. That I am, *expresses the substantiale in me*, but the I cannot be determined, that cannot occur other than through accidents (28:511, emphasis added).

(6) From *Metaphysik L2* (1790–1791):

In the concept of I lies substance, it expresses the subject in which all accidents inhere. Substance is a subject that cannot inhere in other things as accident. The *substantiale* is the proper subject (28:590).

The perdurability of the human soul cannot be inferred from the concept of substance. The ancient philosophers inferred that, because the soul cannot perish through division, it will not perish at all. But this is false, for there still remains yet another perishing, namely, when its powers gradually diminish and disappear, until finally they stop altogether and are transformed into zero or into a nothing.—The soul is not material, matter is composite, and not simple, also no part of matter is simple, which is good to note, for the parts of matter must also constantly be material. But the soul is simple, and thus not material (28:591).

(7) From *Metaphysik Dohna* (1792–1793):

Wolf wanted to derive everything from the faculty of cognition, and defined pleasure and displeasure as an act of the faculty of cognition. He also called the faculty of desire a play of representations, thus likewise a modification of the faculty of cognition ... but this is impossible here. Wolf came to this merely from the cited false definition of substance; thus there were powers which all had to be derived from a basic power. So he assumed the power of representation as basic power—etc. But power is nothing but the mere relation of the accidents to the substance (28:674).

... the soul is simple substance (28:684).

(8) From *Metaphysik K2* (early 1790s):

For the survival of the soul is required the survival of its substance and the identity of its personality, i. e., the consciousness of being the same subject that it was. One attempts to prove the first by the following ground: the soul is simple, thus it is indestructible (incorruptible by inner decomposition) and cannot perish in this way. (The parts of matter indeed remain over, but it itself perishes.) Mendelssohn held this proof not to be adequate: he says ... This proof is not stringent. The soul cannot perish through division, but clearly through remission, through remission of powers (just as consciousness has various degrees of clarity, which become ever weaker, e. g., in falling asleep). The extinguishing of the human soul until complete evanescence can therefore be quite easily thought. There will be no leap here, but rather all can go according to the laws of continuity. With one degree of power the soul is there in one time; between this and the moment where it wholly disappears, there are a multitude of moments where the degrees are various. It seems contradictory to this representation that in all alterations in nature the substance perdures and only the accidents change. But here the talk is merely of bodily substances, which we cognize, but with the human soul we cognize nothing perduring, not even the concept of the I, since consciousness occasionally disappears. A principle of perdurability is in bodily substances, but in the soul everything is in flux (28:763).

(9) From *Metaphysik Vigilantius* (1794–1795):

(2) Substance thought with omission of all inhering accidents (i. e., their determinations) is called the *substantiale*. This remainder is a mere concept that has no determination. It is a something, hence is merely thought or is representable, for the *substantiale* cannot be cognized. Nothing can be cognized if one does not have predicates of the object whereby something is cognized, because all cognitions happen only through judgments. But here only the subject remains without a predicate, therefore no relation between the two. There thus remains left only a representation of a something, but of which one does not cognize what it is ...

So thinking, willing, feeling of pleasure and displeasure are predicates of the human soul. If these are left away, and the soul is thought without these inhering items, then something remains left of which one has no concept, a thought without thinking subjects, and this is the *substantiale* (29:1004–5).

(10) From *Reflexion* 6334 (1795):

It appears that, if one admits that the soul is substance, that one also needs to admit permanence as with bodies. But we can recognize absolutely nothing permanent in the soul, as, e. g., heaviness or impenetrability with bodies.—

Thus is the concept of the soul as substance only a concept of a bare category of the subject to distinguish it from its inhering accidents (18:655).

In 1795, Samuel Sömmerring sent Kant a work, *On the Organ of the Soul*, about the anatomy of the brain and the functions of nerves, informing Kant he would dedicate the work to Kant. In response, Kant sent Sömmerring a four page commentary on the work, with permission to publish it. This commentary spoke approvingly of Sömmerring's empirical research and addressed the question of the relation between empirical research on empirical matters concerning the brain and philosophical reflection on the metaphysics of the soul. Here Kant addresses (a few years before his 1798 *Conflict of the Faculties*) the rightful intellectual domains of the medical and philosophical faculties, respectively. Sömmerring published his essay in 1796 along with Kant's commentary. In addition to the published version, found in the *Akademie* edition volume 12, on Kant's correspondence, and in the Cambridge Edition volume on anthropology, history, and education, three drafts of this letter are included in volume 13 of the *Akademie* edition.

From the published version, we read: "By *mind* [*Gemüth*] one understands only the faculty (*animus*) of combining the given representations and of effecting the unity of empirical apperception, not yet the substance (*anima*) in accordance with its nature wholly distinct from material, from which we abstract here" (12:32n). From the drafts (not in the Cambridge Edition), we read the following:

... this question if taken by the letter destroys itself for it could approximately be translated as what sort of spatial relation is supposed to be recognized between a thing that can in no way be an object of outer sense (because it is simple) and the body of a person that is such an object, for a spatial relation can be achieved only between two objects of outer sense—accordingly the question must be presented differently or mean something different, namely, in which (not local but rather) **virtual** relation of **immediate** influence (which we can clearly not explain) the soul and body in humans, or, more narrowly determined, in the brain (where the ends of all sensory organs are found), stand, and which part of the latter is the next organ of the former to receive representations from this *facultas repraesentativa* or oppositely to effect motions in it (*fac: locomotiva*) presupposes that the person (and likewise every animal) has a soul, i. e., a substance in itself distinct from all material which because of the unity of consciousness must be thought of as simple (therefore itself not again as an object of **outer** sense) therefore not something located in space (*localiter*) but instead only (as the understanding thinks it) as **active** (*virtualiter*) without spatial determination (13:401–2).

We pursue not the immediate effect and action of the soul but instead only the appearances of the same. The former would concern the nonsensible substrate of material as the soul itself is (13:407).

What we see, then, is that Kant believes the noumenal soul is a substance, understood in the spare sense of something that has powers that enables it to produce accidents. But Kant also believes we know nothing determinate whatsoever about this substance; it is as such distinct from all of its accidents, for they merely constitute its state (A360). He also believes there is nothing inconsistent about thinking that the powers of this substantial soul could evanesce to the point where it ceased to exist, so that we are not warranted in concluding its permanence toward the ultimate conclusion of the soul's immortality.

### 3. The History of Kant's Use of the Concept of Substance

If the soul as thing in itself is substance but cannot be shown to be permanent, from where might the rationalists derive the conclusion of permanence, and how might they commit a *sophisma figurae dictionis* of a transcendental sort in the process? Here we need to look at the evolution of Kant's account of the *schematized* category of substance. Already in the early 1760s Kant breaks from the rationalists by rejecting the view that a substance, per se, is self-sustaining, arguing from at least his 1763 *Negative Magnitudes* essay onwards that only an extramundane God is an independent substance and creator and sustainer of other substances (2:202; see also R3879, 17:323 [1764–1769]). In his 1766 *Dreams of a Spirit Seer*, Kant nonetheless recognizes that we assume the permanence of all substance because not doing so would have a destabilizing effect on our knowledge of the world. In the same year, in a *Reflexion*, he argues similarly that “The always lasting duration of substances, i. e., the same age of each with the whole world, cannot be proved, as that it must lie at the base of the method of philosophizing” (R4105, 17:416 [1769]). After 1770 and the clear distinction between noumenal and phenomenal substance, however, Kant relates this assumption of permanence to phenomenal substance in particular. He tells us in R5312, from 1776–1778, that “A phenomenon, which is a substratum of another phenomenon, is not therefore substance but only comparatively. In ap-

pearance we cannot recognize something as substance (this is only a concept of apperception), but rather something only appears as the substratum of appearance, to which everything in appearance is attributed" (17:150). Likewise in *Metaphysik L1* he will explain that "we find in bodies substance that we call substances only by analogy" (28:209). Finally, in the First Analogy, Kant will attach to this phenomenal substance, characterized by extension and impenetrability, the assumption of permanence, arguing that, without this assumption of permanence of phenomenal substance, there could be no unity of time. He thus consistently refers to this permanent substance as "substance (phenomenon)" as in these passages: "Therefore in all appearance that which persists is the object itself, i. e., the substance (phaenomenon) ..." (A183/B227); and "This persistence is nothing more than the way in which we represent the existence of things (in appearance)" (A186/B229). What we also can notice in the first passage from the First Analogy is that Kant not only refers to the substance in question as a phenomenal substance but also as an "object itself" (B227). However, going back to the Aesthetic, we see the meaning of this label "thing itself." There Kant drew the transcendental distinction between a thing in itself in "a merely physical sense" and a thing in itself "transcendentally" (B63). Kant is referring to this physical sense of "thing itself" in the First Analogy when he speaks of the phenomenal object itself. Kant accordingly tells us in the *Prolegomena*, for example, that "permanence can never be proved of the concept of a substance as a thing in itself, but only for the purposes of experience. This is sufficiently shown by the First Analogy of Experience" (4:335).

Now that we see where the concept of permanence has its legitimate role—in reference only to phenomenal, empirical substance—we can also see how, for lack of transcendental reflection on the concept of empirical, phenomenal substance and our manner of assuming its permanence, one might indiscriminately apply the phenomenal, empirical concept of substance, including permanence, to a noumenal substance. And this, I argue, is what Kant sees the rationalists doing.

This interpretation of the First Paralogism as containing an amphiboly, where we fail to engage in transcendental reflection about the nature of the object we are applying a concept to, and whether it is phenomenal or not (B317), is also borne out by a *Reflexion* from 1780, where Kant first mentions the Paralogisms. There Kant presents them under this heading: "Concepts of Reflexion (their Amphiboly) which lead to paralogisms" (R5552, 18:218). Kant refers to this confusion re-



garding the merely transcendental use of the category—here of substance—at the start of the Paralogisms, telling us that the conclusions from the I think “can contain a merely transcendental use of the understanding, which excludes every admixture of experience; and of whose progress, after what we have shown above, we can at the start form no advantageous concept” (A348/B406). This is what Kant refers to a page later when he observes that “So much is lacking for us to be able to infer these properties [permanence] from the pure category of substance, that we must rather ground the persistence of a given object on experience if we would apply to that object the empirically useable concept of a substance” (A349).

In conclusion, then, Kant’s First Paralogism involves a *sophisma figurae dictionis*, hinging on a lack of transcendental reflection. While the soul is a substance in the bare sense of something that has powers that enable us to ground accidents, this indeterminate concept is useless when it comes to the sole objective of rational psychology—to establish our immortality.<sup>1</sup>

---

1 Many thanks to Stephen Palmquist for his useful feedback. Thanks also to Stephen for organizing the Kant in Asia international conference in Hong Kong from May 20 to May 23, 2009, where this essay was originally presented, and to Hong Kong Baptist University for sponsoring the conference.

## 7. Kants Logik des Menschen – Duplizität der Subjektivität

*Ulrich Fritz Wodarzik*

Das verpflichtete sowohl als  
das verpflichtende Subjekt  
ist immer der Mensch.  
– Kant, *Tugendlehre* (6:419)

### 1. Einführung in die Thematik

„Die Logik ist selbst Philosophie“,<sup>1</sup> sagt Kant und ist die „Wissenschaft von der Beziehung aller Erkenntnis auf die wesentlichen Zwecke der menschlichen Vernunft (*teleologia rationis humanae*), und der Philosoph ist ... der Gesetzgeber der menschlichen Vernunft“ (3:542). Philosophische Logik ist die Wissenschaft des Denkens, die über die klassische Logik hinausgeht, weil „die Gesetzgebung der menschlichen Vernunft ... zwei Gegenstände, Natur und Freiheit hat ...“ (3:543). Das klassische Bewusstsein versteht sich als Identität mit dem Sein und der Gegensatz zwischen Denkgegenstand und dem denkenden Subjekt ist zweiwertig. Nur durch ein Subjekt ist Erfahrung überhaupt möglich, denn der Grund der Erfahrung liegt nicht in der Erfahrung. In demselben Ich verbirgt sich auch die Rede von den Geboten Gottes als unsere moralischen Pflichten. Es gibt keine Wirklichkeit an sich, sondern nur eine gedachte Wirklichkeit. Das Bewusstsein davon ist ein Reflexionswissen und dessen Maßstäbe davon liegen in uns. Der Mensch ist umgeben von Objekten und Subjekten. Er steht zwischen Natur und Gott und vermittelt zwischen Endlichkeit und Unendlichkeit. Diese sinnlich-übersinnliche Differenz bewegt und strukturiert menschliches Dasein.

Jedes philosophische System ist durch die philosophische Tradition geprägt, worauf Kant auch selbst hinweist: „Schon von den ältesten

---

1 „Wunderlich: dass die Logik kein allgemein hinreichendes Kriterium der Wahrheit angeben kann“ (16:6 [R1567]; Ebd. 50 [R1629]). Alle Zitate werden in moderner Orthographie geschrieben und durch Band, Seitenzahl angegeben.

Zeiten der Philosophie her haben sich Forscher der reinen Vernunft außer den Sinneswesen oder Erscheinungen (*phaenomena*), die die Sinneswelt ausmachen, noch besondere Verstandeswesen (*noumena*), welche eine Verstandeswelt ausmachen sollten, gedacht ...“ (4:314). Durch die Einteilung der seienden Gegenstände in *phaenomena* und *noumena* hat Kant zusammen mit dem sich selbst erkennenden Ich die klassische Logik erweitert. Den Sinn, den ich einer Sache gebe oder geben soll, kann ich nicht erkennen, sondern nur denken. Erkennen kann ich nur das, was der Fall ist. Denken ohne Anschauung ist immer kontrafaktisch.<sup>2</sup> In dem Abschnitt über die *transzendente Deduktion* erläutert Kant, dass uns keine Erkenntnis a priori möglich ist, lediglich von Gegenständen möglicher Erfahrung. Und er erklärt dort in einer Fußnote:

Damit man sich nicht voreiliger Weise an den besorglichen nachteiligen Folgen dieses Satzes stoße, will ich nur in Erinnerung bringen, dass die Kategorien im *Denken* durch die Bedingungen unserer sinnlichen Anschauung nicht eingeschränkt sind, sondern ein unbegrenztes Feld haben, und nur das *Erkennen* dessen, was wir uns denken, das Bestimmen des Objekts, Anschauung bedürfe ... (3:128).

Kant will damit sagen: „Sich einen Gegenstand *denken* und einen Gegenstand *erkennen* ist also nicht einerlei. Zum Erkenntnis gehören nämlich zwei Stücke, erstlich der Begriff, dadurch überhaupt ein Gegenstand gedacht wird und zweitens die Anschauung, dadurch er gegeben wird ...“ (3:116). Zum reinen Denken gehört also nur der Begriff, d. h. der Sinn, durch den ein Gegenstand gedacht wird. Um zu denken und nicht bloß zu fantasieren brauchen wir unbedingt einen Sinn für das Ansichsein, denn „wir haben einen Verstand, der sich *problematisch* weiter erstreckt als“ die „Sphäre der Erscheinungen“ (3:211). Deshalb müssen Begriffe des reinen Denkens logische Funktionen sein, um die *noumena* zu denken. Mit der Theorie der Vernunftschlüsse sprengt Kant de facto die

---

2 „[D]enken kann ich, was ich will, wenn ich mir nur nicht selbst widerspreche, d. i. wenn mein Begriff nur ein möglicher Gedanke ist, ob ich zwar dafür nicht stehen kann, ob im Inbegriff aller Möglichkeiten diesem auch ein Objekt korrespondiere oder nicht. Um einem solchen Begriff aber objektive Gültigkeit ... beizulegen, dazu wird etwas mehr erfordert. Dieses Mehrere aber braucht eben nicht in theoretischen Erkenntnisquellen gesucht werden, es kann auch in praktischen liegen“ (3:17). Die letzten beiden Sätze markieren genau den Übergang von der Seins- zur Sinnthematik. Diesen Übergang kann die zweiwertige klassische Logik nicht leisten, sie behält aber den Status einer Unterlogik in einer mehrwertigen transklassischen Logik.

klassische Logik,<sup>3</sup> weil diese Schlüsse nicht ontologisch gelten, sondern sich total auf das Denken selbst als eine allgemeine Regel beziehen.<sup>4</sup> Das Noumenon negativ verstanden, bedeutet reine Seinshematik, d. h. theoretische Vernunft und das Noumenon positiv verstanden, bedeutet reine Sinnshematik, d. h. praktische Vernunft.

Der alte Kant schreibt im *opus postumum*: „Transzendentalphilosophie ist das philosophische Erkenntnissystem welches a priori alle Gegenstände der reinen Vernunft in einem System notwendig verbunden darstellt. Diese Gegenstände sind Gott, die Welt und der dem Pflichtbegriff unterworfenen Mensch in der Welt“ (21:81). Die leitenden traditionellen Ideen lassen sich prinzipiell durch die metaphysische Trias Welt-Mensch-Gott fassen. Der Mensch ist ein denkendes Lebewesen und befindet sich immer im Spannungsfeld von Erfahrung und Metaphysik, zwischen Triebhaftigkeit und Pflichtgefühl. Drei Problembereiche haben wir vorliegen: Erstens die Welt der Erscheinungen, zweitens das Denken selbst und dessen Formen und Regeln und drittens das autonome Selbst oder die Person.<sup>5</sup>

Meiner Auffassung nach zeigt sich im transzendentalen Idealismus ein Problembereich, der auf eine nicht-Aristotelische Logik verweist.<sup>6</sup> Kant ahnte das und bemerkte, dass „seit dem *Aristoteles*“ die Logik „keinen Schritt rückwärts hat tun dürfen“ aber „daß sie auch bis jetzt keinen Schritt vorwärts hat tun können“, und zwar deswegen, weil sie „geschlossen und vollendet zu sein scheint“ (3:7). Diese versteckte Aufforderung, über die Logik selbst nachzudenken, wurde von Fichte konsequent erkannt<sup>7</sup> und von Hegel und Günther<sup>8</sup> weitergeführt. Klassisch

---

3 Die Verstandesschlüsse sind konstitutiv für die klassische zweiwertige Logik und gelten für alle Wissenschaften und Mathematik.

4 Hier ist der Unterschied zwischen dem Begriff selbst und seinem Gebrauch wichtig. Das allgemeine Vernunftschlussprinzip lautet: „Was unter der Bedingung einer Regel steht, das steht auch unter der Regel selbst.“ Vgl. *Jäsche-Logik* (9:120). Ein gegebener Begriff ist analytisch und ein gedachter ist synthetisch.

5 Kant bringt diese Trias zur Sprache: „Transzendentalphilosophie ist Erkenntnis des Menschen von sich *selbst* der *Welt* und *Gott*“ (21:157). Ich nenne diesen Sachverhalt das Denkmodell der metaphysischen Trinität: Welt, Mensch und Gott.

6 Die klassische zweiwertige Logik des Wahren und des Falschen, die auf Aristoteles zurückgeht und ihre universelle Gültigkeit weiterhin behält, ist in einer mindestens dreiwertigen Logik eine Unterlogik.

7 Wie aus den Entwürfen der Berliner Darstellungen der WL von 1812 zu ersehen ist und bereits in der Jenaer WL ihren Anfang nahm.

existiert zwischen uns als erlebendem Ich und der Wirklichkeit kein Drittes, denn die Aristotelische Sachlogik ist determiniert durch die Disjunktion „entweder oder“ (*tertium non datur*). Die zweiwertige, zeitlose Logik fordert, dass, wenn eine Vorstellung „wahr“ ist, sie auch prädikativ identisch mit dem Vorgestellten ist. Ist sie das nicht, dann ist sie „falsch“. Sein und Denken stehen sich in der Aristotelischen Logik wie Objekt und Subjekt monothematisch gegenüber. Die antike Logik, axiomatisiert durch Identität, Widerspruchsverbot und *tertium non datur*, spiegelt unser urphänomenales Erkennen wieder. Diese Logik, durch starre Denkgeregeln fixiert auf die sich die Dogmatiker<sup>9</sup> stützen, beherrscht unser gesamtes theoretisches Wissen und klammert das reflektierende Subjekt aus. Deshalb ist diese Logik in den naturwissenschaftlich-technischen Gebieten, jenseits historisch-spekulativer, politischer, religiöser oder sittlicher Fragestellungen, so erfolgreich. Ästhetischer oder religiöser Sinn oder Wert lässt sich nicht durch „wahr“ oder „falsch“ charakterisieren, denn die menschliche Wirklichkeit ist widerspruchsvoll und durch die verschiedensten Kontexte bedingt. Ferner führt die monothematische Zweiwertigkeit in ihrem logischen Gebrauch zum antagonistischen Denken.

Im Zentrum dieses Essays steht die Reflexion oder Selbstbezüglichkeit der Subjektivität, die sich durch ein „zweifaches Ich“ offenbart. Zunächst ist der Mensch im System der Natur als *homo phaenomenon* anzusehen. Darüberhinaus ist er eine Person im Sinne der moralisch-religiösen Vernunft, d. h. als ein *homo noumenon* manifest. Diese Duplizität der Subjektivität bildet den philosophischen Kern des deutschen Idealismus und ist die Dialektik der Subjektivität. Die Unterscheidung zwischen *homo phaenomenon* und *homo noumenon* führte Kant viele Jahre nach der *Kritik der reinen Vernunft* in seiner 1797 geschriebenen *Metaphysik der Sitten* ein.<sup>10</sup> Die Selbstbezüglichkeit und die transzendente Struktur der

---

8 Vgl. Gotthard Günther, I.: *Idee und Grundriß einer nicht-Aristotelischen Logik* (Hamburg: Meiner-Verlag, 1959); II. *Beiträge zur Grundlegung einer operationsfähigen Dialektik* (Hamburg: Meiner-Verlag, 1976 [Bd. 1], 1979 [Bd. 2], 1980 [Bd. 3]); III. *Das Bewusstsein der Maschinen. Eine Metaphysik der Kybernetik* (Agis-Verlag Baden-Baden: Hg. E. v. Goldammer und J. Paul, 2002).

9 Die Dogmatiker haben den „Einfall, sich hinter die allgemeine Logik zu verschansen, und den Schatten des Stagiriten zu beschwören“. Vgl. *Fichtes Werke*, Bd. I, hrsg. v. I. Hermann Fichte (Berlin: Walter de Gruyter, 1971), 499. Im Folgenden mit FW Band, Seite angegeben.

10 Vgl. *Tugendlehre, Ethische Elementarlehre, 1. Teil. Von den Pflichten gegen sich selbst überhaupt*, (6:417–42). Der tiefe Grund für diese Unterscheidung liegt ferner in

Subjektivität erzwingen eine Duplizität der Subjektivität. Daher unterscheidet Kant bereits in der *Kritik* streng zwischen transzendentelem und empirischem Ich. Er thematisierte das ehrgeizig noch im fortgeschrittenen Alter von 67 Jahren in der Preisfrage von 1791: „Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnizens und Wolfs Zeiten in Deutschland gemacht hat“? Die Entwürfe sind als drei unvollständige Handschriften 1804 postum erschienen.<sup>11</sup> Kant hat sich intensiv mit dem Problem der Subjektivität, wie die Paralogismen der reinen Vernunft und die Deduktion der reinen Verstandesbegriffe belegen, beschäftigt.

Der dänische theoretische Physiker und Philosoph Niels Bohr bringt in seinem Buch *Atomphysik und menschliche Erkenntnis* die hier behandelte Thematik auf den Punkt, indem er schreibt: „Hier haben wir es mit komplementären Zusammenhängen betreffend der menschlichen Situation zu tun, die auf unvergessliche Weise in der alten chinesischen Philosophie Ausdruck finden, die uns einschärft, dass wir selbst im großen Drama des Daseins sowohl Schauspieler als Zuschauer sind.“<sup>12</sup> Das zweiwertige dogmatisch-monologische Denken, das keinen Unterschied zwischen Akteur und Spektator macht, versagt, wenn es um das ganze Menschsein geht. „Es ist ganz offensichtlich, daß es gerade diese Unterscheidung ist, in der sich das logische Subjekt des Denkens konstituiert – und sich sowohl von der Welt wie von seinen Gedanken als Drittes absetzt.“<sup>13</sup>

## 2. Duplizität der Subjektivität als Faktum der transzendentalen Differenz

In dem Kapitel *Von der Amphibolie der Reflexionsbegriffe durch die Verwechslung des empirischen Verstandesgebrauchs mit dem Transzendentalen* weist Kant darauf hin, dass sich das Subjekt maskieren muss, wenn es sich selbst als Objekt zum Thema machen will. Das ist die subjektive Zirkularität

---

der Einschränkung der Anmaßung der Sinnlichkeit, vgl. (3:211). Ideen sind *Noumena*, d. h. Vernunftbegriffe oder Gedankenwesen, denen kein Gegenstand in der Erfahrung entspricht.

11 Vgl. (20:255). Auf Seite 270 tauchte der Begriff „zweifaches Ich“ zum ersten Mal auf.

12 N, Bohr, *Atomphysik und menschliche Erkenntnis II, Aufsätze und Vorträge aus den Jahren 1958–1962* (Braunschweig: Vieweg-Verlag, 1966), 15.

13 Günther, *Beiträge* 2, 85. Anm. 8.

(3:265) im transzendentalen Gewand, die im Rahmen der klassischen Logik zu Widersprüchen führt, so dass Kant später vom „zweifachen Ich“ spricht (20:270):

Ich bin mir meiner selbst bewusst, ist ein Gedanke, der schon ein zweifaches Ich enthält, das Ich als Subjekt, und das Ich als Objekt. Wie es möglich sei, dass ich, der ich denke, mir selber ein Gegenstand (der Anschauung) sein, und so mich von mir selbst unterscheiden könne, ist schlechterdings unmöglich zu erklären, obwohl es ein unbezweifeltes Faktum ist ... Es wird dadurch aber nicht eine doppelte Persönlichkeit gemeint, sondern nur Ich, der ich denke und anschau, ist eine Person, das Ich aber des Objektes, was von mir angeschaut wird, ist, gleich anderen Gegenständen außer mir, die Sache. Von dem Ich in der erstern Bedeutung (dem Subjekt der Apperzeption), dem logischen Ich, als Vorstellung a priori, ist schlechterdings nichts weiter zu erkennen möglich, was es für ein Wesen, und von welcher Naturbeschaffenheit es sei ... Das Ich aber in der zweiten Bedeutung, das psychologische Ich, als empirisches Bewusstsein, ist mannigfaltiger Erkenntnis fähig ...

Kant spricht vom Ich in der ersteren Bedeutung, dem logischen Ich, es sei bezeichnet mit Ich<sub>1</sub> (*homo noumenon*). Ferner spricht er vom Ich in der zweiten Bedeutung, dem empirischen Subjekt der Perzeption (*homo phaenomenon*). Dieses sei mit Ich<sub>2</sub> bezeichnet. Die Subjektivität differenziert sich also in den Spektator (spekulierender Philosoph) und in den Aktor, der als Urheber seiner Handlungen fungiert.<sup>14</sup> Die „synthetische Apperzeption“ ist das Ich in der ersten Bedeutung (Ich<sub>1</sub>) und das Ich in der zweiten Bedeutung ist das empirische oder phänomenale Ich (Ich<sub>2</sub>). Ich kann mich von mir selbst unterscheiden und mir dadurch Pflichten oder anderes aufgeben. Nicht vergessen dürfen wir die Welt der Erscheinungen (*phaenomena*), die ich mit der dritten Person „Es“ bezeichne. Damit zeigt sich eine logisch-metaphysische Trinität, die hier mit Es, Ich<sub>2</sub> und Ich<sub>1</sub> bezeichnet sei. In derselben Reihenfolge übersetze ich das mit Welt (*on*), denkend-anschauer Mensch (*Seele, zoe*) und Gott (*Geist, Ideen, nous*).<sup>15</sup> Diese Trinität ist isomorph zu der Fragetrias: „Was kann ich wissen?“, „Was soll ich tun?“ und „Was darf ich glauben?“

14 Aktor (Ich<sub>2</sub>) und Spektator (Ich<sub>1</sub>), d.h. Handelnder und Beobachter sind sprachliche Versuche das Ich als Erscheinung (*homo phaenomenon*) und das Ich als Ich an sich selbst (*homo noumenon*) auszudrücken. Kant führt den Begriff *homo noumenon* deshalb ein, um die Anmaßung der Sinnlichkeit zurückzuweisen und die Pflicht gegen sich selbst plausibel zu machen (vgl. *Tugendlehre*).

15 Vgl. U. F. Wodarzik, „Über die metaphysische Trinität Welt, Mensch und Gott“, *Akten des X. Kant-Kongresses* (Sao Paulo 2005), Bd. 2, hrsg. V. Rohden et al. (Berlin: Walter de Gruyter, 2008), 817–27.

Das Problem besteht nun darin, wie man die beiden Ich-Versionen in einer personellen Einheit ohne Widerspruch zu denken hat.<sup>16</sup> Das empirische Ich, d. h. der lebende Mensch und das logische Ich konstituieren die eine Person.<sup>17</sup> Wie ich, der ich denke, mir selbst als ein Gedachtes sein kann und demnach mich von mir selbst unterscheiden kann, drückt die Aporie der transzendentalen Differenz aus. Das logische Ich<sub>1</sub>, also das was sein eigener Grund ist als *ens causa sui*, d. h. das zeitlose Ich an sich selbst als *homo noumenon*, ist von dem Ich, das da denkt, dem *homo phaenomenon* zu unterscheiden. Die Möglichkeit einer inneren Erfahrung des „Ich denke“ kann nicht vollständig objektiviert werden und ist daher keine empirische Erkenntnis, sondern muss als Erkenntnis des Empirischen überhaupt angesehen werden (3:263) und gehört zu den transzendentalen Postulaten in Form der beiden Erkenntnisstämme Empirie und Ratio. Hier ist der Unterschied zwischen Verstandes- und Vernunftbegriffen maßgebend. „Vernunftbegriffe dienen zum *Begreifen*, wie Verstandesbegriffe zum *Verstehen* der Wahrnehmungen“ (3:244).

Das Faktum der transzendentalen Differenz bezieht Kant auf das Problem der Vermittlung im Sinne des Verstandesvermögen und der unmittelbaren Sinnlichkeit. Er besteht darauf, trotz der Aporie der transzendentalen Differenz, von einer Einheit der Persönlichkeit zu sprechen, obwohl über die Essenz und Existenz des transzendentalen Ich (als Idee, als *homo noumenon*) nichts bekannt ist. Er behauptet, dass ich

mir meiner selbst in der transzendentalen Synthesis des Mannigfaltigen der Vorstellungen überhaupt, mithin in der synthetischen ursprünglichen Einheit der Apperzeption, bewusst [bin], nicht wie ich mir erscheine, noch wie ich an mir selbst bin, sondern nur *dass* ich bin. *Diese Vorstellung ist ein Denken*, nicht ein *Anschauen* [...] und ich habe also demnach keine *Erkenntnis* von mir *wie ich bin*, sondern bloß, wie ich mir selbst *erscheine* (3:123).

Drei metaphysische Komponenten liegen nach dem Bisherigen vor:

1. Ich<sub>1</sub>-Subjektivität (*homo noumenon*) ist der Spektator: ich habe ein Gewissen, bin eine autonome Person und trage die Menschheit in meiner Verantwortung durch die selbstgesetzte Pflicht in mir. Das logische Ich ist überontologisch, nie Gegenstand der Erfahrung sondern bloß reine Idee.

---

16 Ich<sub>2</sub> und Ich<sub>1</sub> kann als ein komplementäres Paar aufgefasst werden; beide Ichs sind verschieden und bilden doch ein Ganzes, nämlich den Menschen.

17 Wie der Mensch und sein Schatten. „Nur will ich anmerken, dass in Ansehung an den inneren Sinn das doppelte Ich im Bewusstsein meiner selbst, nämlich das der inneren sinnlichen Anschauung und das des denkenden Subjekts, vielen scheint zwei Subjekte in einer Person vorauszusetzen“ (20:268).



Von diesem Ich aus erfolgt die Kausalität durch Freiheit. Kant rasoniert: „wir haben kein anderes Richtmaß unserer Handlungen, als das Verhalten dieses göttlichen Menschen in uns, womit wir uns vergleichen, beurteilen, und dadurch uns bessern, obgleich wir es niemals erreichen können“ (3:384).<sup>18</sup> Die einheitliche Synthesis des Mannigfaltigen durch die Denkkategorien, d. h. die synthetische Apperzeption ist das logische Ich in der ersten Bedeutung.

2. Ich<sub>2</sub>-Subjektivität (*homo phaenomenon*) ist der Akteur: Ich bin gewissenlos handelnd als empirisches Ich und Urheber meiner Handlungen durch Sinnlichkeit und Verstand. Dieses empirische Ich ist mannigfaltiger äußerer und innerer Erfahrung fähig. Alle Natur- und Humanwissenschaften werden von diesem empirischen Bewusstsein fundiert. Es ist als rationales Naturwesen der Ursache-Wirkung Relation in der Zeit unterworfen.

3. Es – die Welt (*phaenomena*): Das Gegebene in Form der Erscheinungen.

Es ist das große Verdienst Kants als erster auf die aporetische Verfasstheit der menschlichen Subjektivität hingewiesen zu haben, nämlich die transzendente Differenz in Form der Verdoppelung der Subjektivität.

### 3. Analyse des zweifachen Ich: das Ich-Du Anerkennungsprinzip

Das zweifache Ich von Kant ist Grund für das logische Problem der Mehrwertigkeit der Subjektivität. Das Denken „als geschlossene Totalität und konkrete Realität“ erschöpft sich nicht im Sein der Dinge, „weil alle Subjektivität *im Sein immer nur ein Verhältnis zu sich selbst gewinnen will*. Das

---

18 Der Pflichtbegriff ist hier verhüllt maßgebend: „Gottes Allgegenwart ... zeigt sich uns in dem Gefühl der Pflicht als dem eigentlichen Grundbewußtsein ‚der Gegenwart der Gottheit im Menschen‘: die gebietende Macht des unendlichen Willens als eines uns ‚innigst gegenwärtigen (*omnipraesentissimus*)‘ Wesens vernichtet nicht, sondern sie fordert gerade die Freiheit und Selbstheit der endlichen Willenssubjekte. Kant zielt auf eine Fassung der göttlichen Welttranszendenz und zugleich-immanenz, die gleich weit entfernt ist von der Äußerlichkeit des deistischen Welturmachers wie von den Verschmelzungsneigungen des Pantheismus“. Vgl. H. Heimsoeth, *Metaphysik der Neuzeit* (München: R. Oldenbourg, 1929), 104.

bedeutet metaphysisch Reflexion“.<sup>19</sup> In der klassischen Logik existieren zwei primordiale metaphysische Komponenten, die dinghafte Seinsidentität und die nichtdinghafte Subjektivität. Die traditionelle zweiwertige Logik stellt dem Subjekt (Ich) nur die Objekte (Es) gegenüber. Andere Subjekte in Form der Du-Subjektivität können in dieser Logik nicht thematisiert werden. Bisher wurden alle Subjektivitätsformen, die sich in anderen Menschen verbergen, in der einen idealen Subjektivität schlechthin aufgehoben. Ich erlebe mich als eine innerlich subjektive und unantastbare Existenz (Ich<sub>1</sub>), einen anderen dagegen immer zunächst als eine objektive und physische Existenz in der Welt. Über den Anderen sind mir nur objektive physikalische Erkenntnisse zugänglich, z. B. über die Sprache und verschiedensten Wahrnehmungen, nicht aber seine subjektiven Bekenntnisse. Ein Du wird der andere erst dann, wenn ich ihn als eine Person anerkenne. Die Du-Subjektivität ist demnach nur eine gedachte Subjektivität. Zum methodischen Vorgehen in Ansehung an die Duplizität der Subjektivität ist es von großer Bedeutung, dass durch Abtrennung eines Bereichs der Subjektivität (Ich → Ich<sub>2</sub> + Ich<sub>1</sub>) eine dritte metaphysische Dimension neben der Objekt-Subjekt-Dualität freigelegt wird. Der abgetrennte Subjektivitätsbereich ist die Information oder Kommunikation produzierende Verstandesleistung (Reflexion des Ich<sub>2</sub>), die keinen gemeinsamen Durchschnitt mit der introszenten reinen Innerlichkeit Ich<sub>1</sub> besitzt. Die klassische Idee der Objektivität bleibt dabei unangetastet.

Zum Natursein gehören neben den dinglichen Gegenständen (*phaenomena*) auch andere menschliche Subjekte (*homo phaenomena*). Die zweiwertige Logik des Aristoteles „ignoriert die nicht abzuleugnende Tatsache, daß der Begriff des Nicht-Ich *zweideutig* ist. Nicht-Ich ist erstens: das Du und zweitens: das Ding“.<sup>20</sup> Der Andere begegnet mir als ein Pseudo-Objekt und seine (von mir) gedachte innere Subjektivität ist mir unzugänglich. Es ist meine Option, ob ich den anderen bloß als Ding oder als würdevollen Menschen anerkenne. Jede andere, mir fremde Subjektivität, die mir bloß als Sache erscheint, muss durch einen kognitiven Akt als eine Person (als ein Du<sup>21</sup>) anerkannt werden. Die Du-

19 Anm. 8, *Beiträge* 1, 63,

20 Vgl. Günther, *Idee und Grundriß*, 66.

21 In der englischen Sprache kommt das Du feierlicher und ernster zum Ausdruck als in der deutschen: „I-Thou“ im Gegensatz zu „I-Object“ (I-You), vgl. Donald M. MacKay, „The Use of Behavioural Language to refer to mechanical Processes“, *The British Journal for the Philosophy of Sciences* XIII (1962), 89–103.

Subjektivität taucht wohl zuerst bei Fichte in seiner Erläuterung des Personenbegriffs auf, wie aus der der *Zweiten Einleitung* von 1797 zu ersehen ist: „Der Begriff des Du entsteht aus der Vereinigung des Es und des Ich. Der Begriff des Ich in diesem Gegensatz, also der Begriff des Individuums, ist die Synthesis des Ich mit sich selbst.“<sup>22</sup>

Das Urmotiv für eine transklassische Logik ist das doppelte Subjektsein gegenüber dem Sein an sich selbst, das als Erscheinung erkannt wird. Im Gegensatz zur klassischen Logik wird in der transklassischen Logik die Subjektivität durch eine Ich- und eine Du-Komponente, nicht existentiell, sondern bloß reflektorisch repräsentiert. Die objektiven Dinge sind für jedes Subjekt irreflexiv. Das selbstbewusste reine Denken ist nur vom reflektierenden Subjekt abhängig. Reines Denken ist immer eigenes Denken, daher hat jeder Mensch sein Bild oder seine Wertvorstellungen der Gegenstände, Zustände und Prozesse. „Das Denken denkt sich selber und macht damit seinen eigenen Reflexionsprozeß zum „zweiten“ Gegenstand“.<sup>23</sup> Der „zweite“ Gegenstand „ist ein Bild der Reflexion, die sich selber auf der Gegenstandsebene spiegelt“. Dieses Bild ist von einem ichhaften Denkprozess (Aktor, Ich<sub>2</sub>) produziert und nur für dieses Ich (Spektator, Ich<sub>1</sub>) präsent. Denkende Subjektivität ist immer die eigene und die gedachte immer das Fremde. Fremdheit ist das generelle Kriterium von Gedachtem überhaupt, unabhängig davon, ob das Denkobjekt als irreflexives Sein oder als bewusste Reflexion interpretiert wird.

Wir erleben das Du nämlich weder als Ich – ich bin nämlich einzig und allein für mich selbst Ich – noch erleben wir es als Ding von gleichem Status wie die anderen unbelebten Dinge. Das Du ist auch nicht ein ichhaftes Objekt ... Es ist vielmehr ein Drittes, das aus der zweiwertigen Struktur des Denkens prinzipiell ausgeschlossen ist.

Das Du drängt sich uns auf, weil „es uns mit selbstständigen Denkvollzügen begegnet, die wir in unserer eigenen Reflexion parieren müssen“, daher „zählt es ebenfalls als ein Grund und Motiv unseres Denkens.“<sup>24</sup> Jedes Ich muss seine Subjektivität durch Anerkennung in andere setzen, wodurch es die „Einheit seiner selbst in seinem Anderssein“ erfährt und ist daher immer auf ein anderes Selbst angewiesen. Also liegt die Wahrheit in

---

22 Fichte, I, 502.

23 Vgl. Günther, *Idee und Grundriß*, 335.

24 Günther, *Idee und Grundriß*, 277.

der Anerkennung<sup>25</sup> der Subjekte untereinander, wodurch erst ein Diskurs und eine intersubjektive Sittenlehre möglich werden. Die Abwesenheit der symmetrischen Anerkennung ist die logische Ursache des moralischen Übels, daher bildet das Anerkennungsprinzip die Grundlage einer intersubjektiven Sittenlehre. Die innere Beobachtung meiner selbst ist logisch gleich der Beobachtung anderer Subjekte, die ich mir als Pseudoobjekte, d.h. als Du-Subjekte vorstelle. Ich-Subjektivität und Du-Subjektivität können sich nur physisch in der gemeinsamen Umwelt begegnen, niemals jedoch in Form von intrasubjektiven Erkenntnis- oder Bekenntnisprozessen. Der innergeistige Prozess der freien Willens- und Entscheidungsbildung oder der seelischen Motivation beim Du sind dem Ich nicht zugänglich und umgekehrt.

Wie ist nun aber die Anerkennung einer anderen Subjektivität möglich, wenn die gegenseitigen Bewusstseinsinhalte und -gehalte von Ich- und Du-Subjektivität prinzipiell unzugänglich sind? Diese Frage beantwortet sich dergestalt, indem die Du-Erfahrung, die ein Ich macht darauf beruht, dass „das Ich in der Selbst-Reflexion einen Akt vollzieht, in dem es die Fremd-Reflexion ... als fremde Selbst-Reflexion *anerkennt*“.<sup>26</sup> Dieses Anerkennen zwischen der Ich- und Du-Subjektivität ist zwingend, weil das Du bzw. das Ich diese Anerkennung, wenn Wahrheit oder das Ganze im Vordergrund stehen soll, bestätigt haben will. Denn wenn „das Ich die subjektive Selbst-Gewißheit seines Denkens nie auf das Du übertragen kann und von dem Du dasselbe gilt, dann erstreckt sich diese Unübertragbarkeit auch auf jenes Moment der Wahrheit, das als Erlebnisevidenz an die private Introspektion des isolierten Subjektes angeschlossen ist“.<sup>27</sup>

Die Verdoppelung der Subjektivität ermöglicht eine logische Thematisierung der Du („Thou“)-Subjektivität. Ohne ein Du kann Ich mir logisch niemals selbst gewiss sein. Das Ich-Du Verhältnis ist bei Kant nirgends expliziert, aber er notiert in der *Kritik* unter dem Abschnitt *Von den Paralogismen der reinen Vernunft*, dass es

befremdlich [erscheint], dass die Bedingung, unter der ich überhaupt denke, und die mithin bloß eine Beschaffenheit meines Subjekts ist, zugleich für alles, was denkt, gültig sein solle, ...: daß alles, was denkt, so beschaffen sei, als

25 Der Anerkennungs-begriff stammt von Fichte, vgl. *Grundlage des Naturrechts* 1796 und *Vernunftrecht* 1812. Die Anerkennung ist ein Trieb wie der Naturtrieb, er kann von uns unterdrückt werden oder nicht.

26 Anm. 8: Günther, *Bewusstsein der Maschinen*, 163.

27 Ebd., 166.

der Ausspruch des Selbstbewusstseins es an mir aussagt. Die Ursache aber hiervon liegt darin: daß wir den Dingen a priori alle die Eigenschaften notwendig beilegen müssen, die die Bedingungen ausmachen, unter welchen wir sie allein denken. Nun kann ich von einem denkenden Wesen durch keine äußere Erfahrung, sondern bloß durch das Selbstbewusstsein die mindeste Vorstellung haben. Also sind dergleichen Gegenstände nichts weiter, als die Übertragung dieses meines Bewußtseins auf andere Dinge, welche nur dadurch als denkende Wesen vorgestellt werden (3:265–6).

Kant bemerkt, dass die Moral mit der Idee der Menschheit zusammen-gedacht werden soll und ferner: „Der Mensch ist zwar unheilig genug, aber die Menschheit in seiner Person muss ihm heilig sein“ (5:87). Das ist die Gegenüberstellung von Mensch (*homo phaenomenon*) und Person (*homo noumenon*). Und noch wesentlicher: „Handle so, daß du die Menschheit, sowohl in deiner Person, als in der Person eines jeden andern, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchest“ (4:429). Das kann nur durch einen Ich-Du Dialog in Form eines transzendenten Intersubjektivitätsprinzips möglich werden. Das Problem des absoluten Subjekts, d. h. ohne Ich-Du Relation, liegt nicht darin, dass es alles auf sich bezieht, sondern darin, dass das absolute Ich sich selbst nicht feststellen und festhalten kann.<sup>28</sup> Die Ich-Philosophie ohne den Du-Bezug kreist unverstanden im Subjektivitätszirkel und löst sich selbst auf. Nach Natur und Denkgesetzen allein ist es völlig unbegreiflich eine andere Person als eine von mir unabhängige Persönlichkeit zu erkennen. Das andere Individuum ist für mich zunächst ein Es, dann ein empirisches Ich und schließlich erkenne ich es als eine logische Person an. Die Aporie des zweifachen Ichs bei Kant findet seine Auflösung, wenn wir die „doppelte Subjektivität“ in Form einer Ich-Du Relation per Anerkennung deuten.

Das menschliche Ich teilt sich in ein profanes Naturwesen (*homo phaenomenon*)<sup>29</sup> und in ein heiliges Vernunftwesen (*homo noumenon*) auf. Nur durch Anerkennung setze ich in ein anderes Subjekt, das für mich immer ein Du ist, d. h. sein soll, ein logisches Ich. Die Ich-Subjektivität und die Du-Subjektivität unterliegen einer symmetrischen Relation als ein Umtauschverhältnis. Die Vernunft bezieht sich vermittels des Verstandes auf einen sinnlichen Gegenstand, d. h. auf die Erscheinung (*phaenomenon*). Der Verstand sagt mir, „da ist ein individueller Gegenstand, ein Mensch“ (*homo phaenomenon*) und die praktische Vernunft, „da

28 Vgl. W. Schulz, *Fichte, Kierkegaard* (Stuttgart: Neske, 1977), 22. W. Schulz hat sich in all seinen Werken ausführlich mit der Ichverlorenheit auseinander gesetzt.

29 „Der Mensch ist selbst Erscheinung“ (3:374).

ist eine freie, autonome und unbedingte Person“ (*homo noumenon*), wie ich eine unbedingte Person bin.<sup>30</sup> Es ist das zweifache Ich im Bewusstsein meiner selbst, das diesen intersubjektiven kognitiven Akt leistet. Vernunft kann sich relational in ihren Gebrauch auf dreierlei richten: 1. Es – Ich<sub>2</sub>, d. h. meine erscheinende Umwelt und ich, d. h. reine Seinshematik, 2. Ich<sub>2</sub> – Ich<sub>1</sub>, d. h. ich und ich selbst und 3. Ich<sub>1</sub> – Es, d. h. ich selbst und Erscheinung.

Wie erkenne ich beim anderen so was wie eine menschliche Seele, Würde oder Persönlichkeit? Wie mache ich ein außerhalb meiner Subjektivität liegendes, mir fremdes Objekt zur Persönlichkeit? Was für eine Erfahrung oder Erkenntnis mache ich da? Hier liegt scheinbar ein Vermittlungsproblem vor. Persönlichkeit ist ein Doppeltes im Sinne einer Ich-Du Beziehung, die sich auf eine Ich-Wir Form leicht erweitert und daher eine menschliche Gemeinschaft möglich macht. Ich weiß, dass ich ein Teil der Welt bin, d. h. neben den Dingen der Umwelt und den mir fremden Subjekten der Mitwelt, denn: „Alle [meine] Vorstellungen haben eine notwendige Beziehung auf ein *mögliches* empirisches Bewusstsein ... Alles empirische Bewusstsein hat aber eine notwendige Beziehung auf ein transzendentes Bewusstsein, nämlich das Bewusstsein meiner selbst, als die ursprüngliche Apperzeption“ (4:87).

Im Selbstbewusstsein weiß ein Ich sich selbst als ein Ich im Sinne der Einheit des Wissenden und Gewussten. Der Antagonismus zwischen Welt und Ich oder Sinnlichkeit und sittliche Pflicht, d. h. der Unterschied zwischen diesem Menschen hier und jetzt und der Menschheit kann nur durch ein gewissermaßen höher geordnetes Drittes umfasst und vermittelt werden. Dieses Dritte ist der denkende und lebendige Mensch. Der Mensch findet sich in seiner Endlichkeit immer schon vorausgesetzt, er ist immer schon in der Welt und weiß sich als ein Naturwesen, ein Geschöpf inmitten der Schöpfung. Dass diese Endlichkeit als Sinn oder Wert gefasst werden kann, ist nur möglich auf Grund der synthetischen Einheit seiner Person. Ein Mensch ist unmittelbar gegeben und wird durch kognitive Vermittlung eine moralische, freiheitliche und rechtliche Person. Von selbst ist er das nicht. Die Ich-Du Dialektik ersetzt jede Spielart des Ich-Monismus, weil die Wirklichkeit des Menschen radikal und unauflöslich den Gegensatz zwischen einem Ich und einem Du konstituiert. Die Ich-Du Relation wurde von L. Feuerbach, M. Buber, K. Heim, J. Cullberg und G. Günther weiter entwickelt und begründet mögliche intersubjektive soziale Räume, die logisch begriffen werden können.

---

30 Vgl. auch: D. Sturma, *Philosophie des Geistes* (Leipzig: Reclam, 2005), 65.

#### 4. Schlussbemerkung

Das Wesen der transklassischen Logik erkennt man gut an dem folgenden Gespräch: „Als Platon von dem Menschen sprach, sagte Diogenes, der Kyniker: ‚Ich sehe wohl einen Menschen, aber nicht die Menschheit.‘ ‚Richtig‘, erwiderte Platon; ‚denn Augen, womit man den Menschen sieht, hast du wohl, aber den Geist, womit man die Menschheit sieht, den hast du nicht‘“. <sup>31</sup>Den Menschen sehen wir in der Welt der Erscheinungen mit den Augen, aber die Menschheit, d. h. seine Persönlichkeit „sehen“ wir nur mit dem Geist.

Ausgangspunkt dieser Untersuchung ist das „zweifache Ich“ als Faktum der transzendentalen Differenz, das sich in eine Ich-Subjektivität und eine Du-Subjektivität distribuiert. Die daraus folgende Reflexionslogik erzwingt eine metaphysisch-triadische Grundstruktur. Das klassische Paar phänomenon und homo noumenon transformiert sich in der transklassischen Logik in die Triade: phänomenon, homo phänomenon und homo noumenon. Die logische Tradition irrt, wenn sie annimmt, dass die klassische zweiwertige Logik das Denken erschöpft.

---

31 Von mir paraphrasiertes Hegelzitat. Vgl. G. W. F. Hegel, *Werke 19*, Suhrkamp, 3. Aufl. (Frankfurt am Main: Suhrkamp, 1998), 39.

## 8. Antinomy of Identity

*Michael Thompson*

### 1. Questions of Identity

A principle of individuation, the self (personal, human identity), other, memory, substance, substrate, particularity, universality, contingency, necessity, essence, action, organization, disposition, body, consciousness—each of these heavy, meaning-laden metaphysical concepts plays a role in the historical development of the idea of identity. The chief difficulty in examining theories of identity is the inconsistent but related use of these terms amongst canonical authors from the history of philosophy. Oversimplified, we can separate the history into two factions: the proponents of personal identity and its detractors. To continue this simplification, I wish to suggest this division roughly follows the rationalist and empiricist divisions within philosophy. Canonical masters like Descartes, Leibniz, and Kant argue convincingly for a notion of identity that is clarified by an act of consciousness perceiving its own intuiting. From the empiricist camp we find the mighty Hume looming over this debate to assert in his skeptical manner that identity or self is merely an illusion of the imagination.<sup>1</sup> In Nietzsche we find an adamant denial of the self and yet simultaneous affirmation of personal identity. In contemporary literature the debate still rages. Supporting the classical proponents of identity we find figures such as Paul Ricoeur, Christine Korsgaard and, with Nietzschean predilections, Alexander Nehamas. Continuing Hume's doubt are contemporary figures like Derek Parfit and Richard Rorty.

---

1 Two exceptions need to be noted. In the rationalist tradition, Spinoza must finally land with the opponents of identity and claim every person/thing is a mode of the single substance. This very line of inquiry prompts a question about the identity of this one substance. Unfortunately, I will bypass this perhaps illuminating question in favor of a more topical approach. The second exception noted is John Locke. Although in the empiricist tradition, Locke does eventually affirm identity, as a forensic matter.



In Eastern schools of thought we find an interesting parallel. Confucian teachings inform us that who an individual is—that is, their identity—is founded upon the empirical interconnections one finds in their relationships with others. In the *Analects* Confucius teaches that a person's identity is determined by the five great relationships, chief amongst these being the family; it does not concern metaphysical speculation on the possibility or status of a soul. Who one is, according to Confucius, depends upon the roles and relationships one enacts. This approach mirrors the empirical approach found in Western philosophy by elaborating the multifarious ways<sup>2</sup> we can define our personal identity by external, empirical, determinations. Our identity, according to this approach, is a matter of determining where one fits in within the larger context of social relations. Unlike the Western tradition, however, Confucianism does not lay its emphasis on the physical or material constituents of the human body. Rather, the emphasis is placed on the part-whole relationship of individuals (so achieved) and their social settings.

Taoism, on the other hand, mirrors the rationalist approach. According to the Taoist teachings, a single totality of the Tao is all that is truly metaphysically real, and personal identity must be viewed in light of the single unifying principle of reality. Interestingly, the Taoist conception of personal identity is dependent upon a part-whole relationship, but not one determined by social roles. Rather, its identity is understood merely as a manifestation of the Tao. Taoism mirrors the rationalist approach by insisting upon a transcendental ground as the basis for maintaining personal identity, even though personal identity is merely a singular manifestation of the single unifying principle, the Tao. From East to West, it would seem, a trenchant debate about what determines who and what we are—that is, what determines our personal identity—still looms large in both pragmatic and metaphysical discussions of the self.

From this protracted debate concerning identity—more importantly, personal identity—arises what I call the antinomy of identity. Much like the antinomies found in Kant's *Transcendental Dialectic*, both positions present arguments, founded on either a *reductio ad absurdum* or an internal contradiction. Following antinomic form, authors on either side of the debate assume the opposite position of the one they are attempting to prove and show how such an assumption leads to a contra-

---

2 Lao-Tzu, *Tao Te Ching*, trs. Addis and Lombardo (Indianapolis: Hackett, 1993), 4.

diction and thus consequently assert the opposite, their own, as the only viable stance. Proponents of personal identity assume the contingency of the self and show the inconsistencies of such a relativistic view. Opponents of personal identity assume a universal characterization of the self and show the absurdity of proving personal identity on the grounds of a characteristic shared by all. I would like to suggest that the antinomic nature of discussions of identity is based upon ambiguous definitions of the heavily metaphysical concepts enumerated in the opening sentence of this essay. In each argument an outdated and naïve version of these concepts is employed. I contend that, when understood properly, the two opposing camps may be brought into harmony and made consistent with each other. To Kant's dynamic antinomies, I wish to add the antinomy of identity. This is to say, the propositions of both factions are true. It is only within the proper context or framework that we can see the truth of both positions. Moreover, in addition to affirming the truth of both positions, I wish to illustrate how one can synthesize the two positions into a unity that affirms human essence, embraces personal identity, and celebrates contingency.

## 2. The Antinomy

In order best to illustrate the basic arguments in the antinomy of identity, I propose to sketch the argument in terms of thesis and antithesis and the subsequent proofs of each.

### Thesis

Identity consists in having a universal constant that remains unchanged over time. This constant is the unity of consciousness in an individual.

### Antithesis

There is no identity; each "being" is a concrete particular, with no underlying continuity. A particular is constituted by the concrete contingency of a particular situation according to the specification of space and time. Such a "being" changes with the contingency of the situation.

### Proof

Let us assume there is no constant that remains in any particular that grants unity to such a particular over time. This being so, every proposition made concerning the lack of unity of any particular being is asserted by a particular

### Proof

Assume there is a transcendental identity that remains constant and allows for any being to make assertions and propositions. For any such affirming being there must be a continuity to make propositions intelligible. Because this continuity

being that is not the same as, not identical with, the being that comes antecedently in time. Furthermore, any addition, change or modification of such a being creates a new being that cannot be identified with the previous being. Yet, any proposition is composed of particular utterances asserted in a temporal sequence. The time elapsed from the beginning of any proposition to the end of said proposition must be accounted for. Furthermore, any declaration itself constitutes a change within the speaker/affirmer. Thus in the act of any propositional declaration, the being affirming the proposition is not the same at the beginning of the proposition as the being at the end. As no continuity of being obtains from the beginning to the end of the proposition, no such proposition can be made regarding the inconstancy of being. Hence, in order to affirm the impossibility of a constant that remains unchanged over time, one must assume an affirming subject that remains the same. This is transcendental identity, necessary for any propositional affirmation. Some continuity of the affirmer must obtain in order for the affirmation to be possible. In order to underwrite such continuity, there must be some constant that continues through the entire proposition. Without such a continuity, each element of the proposition, because each is discrete and at a different moment in time, will be affirmed by another being and as such will not be a coherent proposition understood by any particular being. Let us call this continuity the unity of consciousness—one that is necessary in order for a proposition to be made intelligibly.

Thus there is identity.

is not wed to any particular being, it is a transcendental requirement for any intelligible proposition, and it can be affirmed that it is the formal requirement necessary for any assertion whatsoever. Owing to the formal nature of this transcendental requirement, it is not specific to any particular and we can affirm it is formally the same in any particular affirming being. But, because of the universal nature of the transcendental requirement, the identity of any particular cannot be ascertained from this criterion. This absolute requirement, common to all affirming beings, does not provide any particular being with identity—unless we wish that all beings are identical to one another based on this shared requirement. Such universal, absolute conditions do not allow for any particular identity.

Thus there is no identity.<sup>3</sup>

---

3 Once again, the specter of Spinoza appears in this formulation. However, we cannot concede that a transcendental requirement for the intelligibility of propositions will provide the conditions to deduce that there is only one substance. The suggestion remains that there may be identity, but only of the one substance, individuated by means other than those transcendental.

### 3. Observations on the Antinomy

As we can see from the antinomy above, the true issue discussed is one concerning personal, human identity and a specification of consciousness that allows an affirmation of individuality. The philosophical issue of numerical identity and matter seems to be settled by the approach found in Locke's *Essay Concerning Human Understanding* when he states that identity concerns the possibility or impossibility for two things to be in the same place at the same moment in time. He concludes that two beings cannot exist in the same place at the same point in time. If such a condition obtains, we have only one thing, one being. Furthermore, we can differentiate between two beings by analyzing their constituent parts. If a being should change its constituent parts, then it has changed and hence is not identical with what it was before.<sup>4</sup> Predating Hume, but continuing the Humean line of argument, Locke's analysis suggests this is sufficient for non-living material beings, but when discussing "a mass of matter and a living body—identity is not applied to the same thing."<sup>5</sup> It would seem that if we were able to reduce human identity to its material constituents, then the antinomy could be solved rather quickly. However, owing to our scientific, or perhaps merely nutritional, information, we know that the constituent particles of our body are constantly changing. Locke concludes, when we are discussing living beings and living things we are discussing more than just aggregates of particles. Thus, for living beings identity must be characterized in terms other than numerical sameness.

The case of Locke also highlights one trenchant problem in the discussion of the antinomy. This problem is the precise nature of this transcendental condition necessary for intelligibility. *Gratis à Descartes*, by Locke's time, philosophy is inundated with substance metaphysics. With Descartes' reformulation of philosophical vocabulary, modern philosophy and personal identity is marked by the dubious debate concerning the nature of incorporeal substance. Following Descartes, one could assert that the transcendental condition necessary for identity in

---

4 John Locke, "Of Identity and Diversity", in *Self and Subjectivity*, ed. Kim Atkins (Maldin, Ma.: Blackwell, 2005), 25.

5 Locke, 26.

living beings is immaterial substance, instantiated in extended space, material substance. However, plausible as this may sound, Descartes' explanation of his dualism has been shown untenable.<sup>6</sup> Yet, opponents of identity do indeed argue against precisely this type of metaphysics. Those who deny a strong sense of identity cite the impossibility of an enduring substance based on our empirical knowledge of how living bodies are constructed and grow. In short, the empirical tradition denies identity by citing the physical conditions necessary for living bodies and denying any non-corporeal substances. Thus substance metaphysics obscures the issue of identity rather than aiding in clarification.

Locke attempts to solve this problem by granting the dynamic nature of physical bodies and accounting for personal identity by appealing to disposition and memory. In so doing, he claims that personal identity amounts to claiming "identity is preserved in that change of substances by the unity of one continued life."<sup>7</sup> Likewise, Confucius argues for an empirical identity based upon the relationships one finds oneself in. With changing circumstances one may find oneself performing different roles in various relationships, but these relationships may be seen as different performances within a complete narrative. While Locke and Confucius do provide a plausible account for identity, by prescribing that humans "own" their memories and roles, and attempt a coherent, continued life-plan, they open the Pandora's box of personal identity by claiming that if one has a complete break from one's life-plan, a break from continuity and narrative, then one may possibly become an entirely other person.<sup>8</sup> In addition, Locke adds to the discussion the metaphysically loaded terms, memory and disposition, while failing to discuss the ontology of these terms in a study of consciousness. In his final analysis, Locke concludes there may be personal identity, but it need not persist through a lifetime.

The sophisticated proponents of identity grant the substantial objection cited above, and contend another source for personal identity. The proponents of the thesis of our antinomy wish to assert a stronger claim than Confucius and Locke. They wish to claim that, although there may be flux in the corporeal nature of our bodies, and even though there is change in the psychical constitution of our mental life, there must be a

---

6 Descartes fallaciously appeals to the pineal gland as the mediator between the two substances.

7 Locke, 27.

8 Locke, 30.

unifying principle that underwrites all our cognitive functions. In other words, there must be an 'I' that thinks. For Kant, the most sophisticated, modern advocate of personal identity, this 'I' that is the absolute subject of all my mental life is precisely the foundation for personal identity. For Lao-Tzu, one finds the universalizing principle not in consciousness, but in the Tao itself, human consciousness being a manifestation of it. Unifying the discrete sensorial perceptions encountered in sensation and the consequent shift in attention that accompanies these perceptions, conjoining the discrete perceptions provided by memory, even merging the disparate dispositions contained in a life-plan is "the absolute unity of the thinking subject."<sup>9</sup> This unity of the thinking subject, a unity of consciousness, is not particular to any specific individual, but is applicable to all beings that think. Moreover, Lao-Tzu may contend the strongest sense of unification, not merely in the consciousness of an individual, but a unification with the cosmic principle Tao. Thus all humans must—that is, transcendently must—have some unifying principle of all their perceptions, and this principle is what we call personal identity, or the self; it may even be seen as a part of a greater identity, the Tao.<sup>10</sup>

Just on this unity Kant and Lao-Tzu are liable to criticism. David Hume's argument contends that such unity is an illusion that complicates the issue. Hume claims this unity is in fact not a unity. We never have an impression of unity<sup>11</sup> but merely feel all our perceptions flow smoothly from one to another in an ordered, principled sequence. A unified relation of perceptions, produced by the faculty of the imagination according to Hume, is a "confusion and mistake, and makes us substitute the notion of identity, instead of that of related objects" that

---

9 Kant, *The Critique of Pure Reason*, tr. Norman Kemp Smith (New York: MacMillan, 1965), A352.

10 This is also the position Christine Korsgaard assumes in "Personal Identity and the Unity of Agency: A Kantian Response to Parfit", in Daniel Kolak and Raymond Martin's *Self and Identity* (New York: MacMillan, 1991). While Korsgaard employs this transcendental method, she ultimately bases the need for unity of consciousness on practical reasons; that is, on the necessity for humans to act in their environment. She argues that motivation, psychological, moral, or existential, provides a unity to all consciousness. In the case of split-brains, there may be other means, even more practical (and bodily) that unite the two halves of a brain with a severed *corpus collasum*.

11 Moreover, without any impression we cannot have an idea, whether true or false.

presumably give rise to it.<sup>12</sup> In other words, we are never conscious of this so-called unity of consciousness. While a Humean critique concerning the continuity of perceptions may be applicable, such a critique does not undermine the greater Kantian point that there must be some “vestibule”, a transcendental requirement, that all the disparate perceptions must flow into in order to have the possibility of any continuous relation of ideas. Being conscious of this consciousness, or a lack of thereof, may indeed be no reason to dismiss the argument entirely.

From the historical development of the concept of identity we see how metaphysically biased terms have added confusion or—dare I say—created the antinomy itself. With the inception of Cartesian metaphysics of substance the stage is set for rancorous argumentation regarding numerical, material, identity and personal identity. Deniers of identity cite the changing nature of physical substance and ask for proof of some non-material substance that endures through time; to such a request the Cartesians (and rationalists in general) are found wanting. But even should we move away from the material debate and move to consciousness, deniers of identity cite the changing nature of perception, the dubious reliability of memory and the possibility of amnesiac or anesthetized cases to illustrate that there need be no unification of experience under the auspices of personal identity. They conclude, with Richard Rorty, that any being is concrete, particular, and changes with stimulation either from environment or propelled by mental states. Because of the contingent nature of each particular and the dynamic relations any particular has with any occurring state at any particular time, either intentional or environmental, there is no enduring unity of self—there is no identity.<sup>13</sup> To the purely formal requirement of the Kantian unity of apperception, such particularists respond that such is an empty concept that provides no meaningful content that enables us to assert any personal identity.

On the other side of the debate, proponents of identity assert that there must be some unification of consciousness, and this is precisely what endures. Although we may not predicate such a unity, one must be asserted, for without such there can be no stream of consciousness, no comparison, no judgment (A116/B142). Furthermore, those who advocate identity make a stronger claim than merely noting the contin-

---

12 Hume. “Of Personal Identity”, in Atkins, 39.

13 A. Rorty, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989), 37.

gency of each individual's experience. They affirm the changing capacity, both physical and mental, inherent in human existence. But in order to register this change, a unifying principle, a unity of consciousness, must be in place. In other words, any being that thinks and deserves the accolade of human, must have some continuity of the self that endures and allows them to enact the cognitive process that registers change. Without such a unity, thinking and intelligibility are impossible.

The debate between personal identity and its deniers attempts to address a landscape fraught with metaphysical pitfalls. One half concludes an absolute, universal, essential self, necessary for human beings to be what they are. The other half cites contingency, particularity, corporeal substance, and memory to conclude no such self exists. The debate eventually devolves into a conflict between essentialists and particularists.

#### 4. Solution to the Antinomy

I propose in this final section to take the good and omit the bad from the above arguments, in order to conclude with what I prefer to call the fact of identity. Because of such acrimony between the two factions, an enterprise of this kind may not be well-advised, lest I make enemies on both sides of the debate. However, I believe such an attempt to be feasible and promising. To side with either camp seems to overlook serious metaphysical truths (perhaps assumptions) about the world we live in. To side with the antithesis of personal identity overlooks what each of us experiences as a self in our daily experiences: the phenomenon of the self. To side with the thesis presents an empty formalism that smacks of metaphysics and esoteric academia, once again overlooking the phenomenon of living in a world. I believe that a phenomenological approach to the self, personal identity, and consciousness may provide an essential scaffolding, a structure of the self, that accounts for how we experience the world.

One such approach, and a very Kantian one at that, is to look to the actions we perform every day. Christine Korsgaard takes this approach and states that a "pragmatic unity is the unity implicit in the standpoint from which you deliberate and choose."<sup>14</sup> In practical situations one is forced to interact with one's situation. Doing so involves choice and, in

---

14 Korsgaard, 324.



order to implement choice actively, it is necessary to deliberate and choose. From a practical standpoint, we must act as if we have a self, a self that grounds our deliberations and choice. Such practical considerations offer a plausible account for why we must think *as if*, but offers little reason to assert the truth of personal identity. Such a regulative ideal follows typical resolutions to antinomic paradoxes, but offers little more than faith in a self. The option I prefer resembles Korsgaard's approach by virtue of the fact that it also recognizes the necessity of choice in a situation, but provides more theoretical support to explain why a self must be posited in order for an agent to make choices.

As alluded to in the beginning of this essay, I believe the antinomy of identity may be resolved when examined in the proper context. It is to this key term, context, that we must turn to find resolution to the conflict. The context we must turn to is the context of the being we have been discussing thus far—the context of living beings, experience. Furthermore, I wish to continue with the implication of the antinomy and reduce the context to those of living, human beings. The possibility of personal identity has been the real underlying issue at stake in the antinomy. Apparently, if we side with either faction we fail to find a substantive, meaningful account of personal identity: with the one we have substance and perhaps meaning, but no identity; with the other we have identity with no substance and no meaning.

A merely formal account of personal identity fails to provide a substantive account of human experience, while contingent accounts of human experience fail to provide identity. To exact both personal identity and meaning from the antinomy of identity we must turn to the human context and uncover the parameters of such. In other words, we need an account of human existence that affirms personal identity and provides a robust account of individuality. The phenomenological investigations led by Martin Heidegger and Maurice Merleau-Ponty provide just such parameters. Heidegger's ontological investigation into the meaning of *Dasein* (crudely defined here as individual humanity) yields that humans as such are temporally conditioned through and through.<sup>15</sup> Unless we can elide time, humans are fundamentally defined by temporality. In the case that humans escape the grips of time we may even say they have transcended humanity and they no longer belong to a study of humanity. From the moment of origination, birth, to the pe-

---

15 M. Heidegger, *Being and Time*, tr. Macquarrie & Robinson (San Francisco: HarperSan Francisco, 1962), 383.

nultimate moment, the one just before death, we are constantly engaged in time. Our existence is temporal. Hence human existence, if marked by no other universal, is demarcated by temporality. In the experience of time, we not only mark the passing of time in consciousness, but, more importantly, we orient ourselves to the future and the past. The present moment, as St. Augustine rightly notes, is the razor's edge between past and future. Furthermore, adds Heidegger, who we are at any given present moment in time is the combination of antecedent history and our projection into the future.<sup>16</sup> Contrary to the deniers of essence, the particularists, this temporal orientation is found in the Dalit of India, the billionaire of America, or the factory worker in China. Temporality marks human existence. It is essential to human nature.

When we examine human experience more closely, we find another characteristic that demarcates human existence: the body. In his investigations of perception, Merleau-Ponty argues convincingly for the body as the "zero point" of orientation in the world.<sup>17</sup> All perceptions begin and end with my body. Sensation obtains in the interaction between myself and world. Perception of such sensation is the internalizing, recognizing, and attending to the data of sense experience. Furthermore, memory is just such a perception, one not attended by sense experience; it is, rather, a process of recollecting past experience or—worst case scenario—a process of recombination fueled by disparate perceptions and combined in fantasy. Despite the difference between physical sensation and mental perception, the origin is the same: the self, the body. On the physical level, the body is the contact point for this individual and the world. At the level of consciousness this body is at first the contact point for this individual and the world and, secondarily, the means enabling me to recall. Like temporality, characterizing human experience as bodily precludes neither the Dalit, the billionaire, nor the laborer. All human experience is characterized, given a sense of style, by the body.

Furthermore, each individual has a unique body. Following the particularists, we may affirm the truth of the contention that each being is composed of discrete masses of matter. Numerically, each individual has identity. But more than just numerical identity, each individual body is

---

16 This very orientation and subjection to time may be what marks the incorporation and transformation of the Tao into the world of ten-thousand objects.

17 Merleau-Ponty, *Phenomenology of Perception*, tr. Colin Smith (New York: Routledge, 2002), 115.

the origin for that being's perceptions. The uniqueness of each individual body suggests that, while the structure of embodied consciousness is the same for each human being, the particularity of each individual can be deduced from its constituent material as well as the impossibility to be in the same place at the same point in time. Merleau-Ponty continues to describe this body-schema in a variety of ways. We are spatial and mobile agents. We, too, are sexual creatures. As a consequence of acting within our world, we are also expressive beings.<sup>18</sup> These characterizations of who we are, according to Merleau-Ponty, are not simply modes of our being, but are, rather, at the root of all our activity in the world. We cannot recuse ourselves from any one of these fundamental characteristics; they shape and define who we are as individuals through a universal structure. Through these universal structures we find ourselves in the world with a particular style unique to each individual. With these phenomenological insights we can thus satisfy some of the particularists' charges of empty formalism to any structuralist approach and provide a principle of individuation.

Yet one facile objection still persists; the charge contends that *all* beings possess some kind of body and are subject to the progression of time.<sup>19</sup> How *any* being may be said to possess personal identity by the criterion outlined above must be addressed. Everything, it might be said, possesses personal identity. What this objection fails to note is the particular way humans orient themselves to time through a collation of bodily orientations in consciousness. Among the various perceptions experienced through the zero point of the body belongs time. But, more importantly, it is the human perception of time, with its particular situational, bodily historicity and its particular projection into the future as concrete plans for this body, that differentiates humans from any other being. Based on our perception of our unique history and our particular projections into the future, humans care about their future in a unique way. Because humans have the capacity to orient themselves to the future in this unique way (through the past), we may say that human temporal orientation is unique to this conscious being. Because the future matters to humans, they must choose their future. Personal identity is

---

18 Cf. *Phenomenology of Perception*, chs.3, 5, 6.

19 Such an observation affirms the ontological commitment of Taoism, by noting that human consciousness is merely one manifestation in the world of the ten-thousand objects, yet fails to note the unique way human consciousness styles its understanding of the world of objects and the Tao.

a synthesis of these contingent, bodily, historically dependent facts and the unity of consciousness necessary to effect an understanding of these facts.

Finally, I would like to make one last connection. I think there is a salient relationship between this ability to choose our future and consequently our present (defined by both our bodily constraints and historicity and through projection into the future) and Nietzsche's prescription to cultivate a sense of style, to choose to be who you are—or better yet, to choose to be today who you will be. Any individual's style will be a combination of those Merleau-Pontian fundamentals of personal body experience in conjunction with that individual's projects. With this final suggestion, I believe that action once again takes a central role in identity. If we choose to be who we are or will be, we de facto enact our identity daily as we affirm original choices or reaffirm those commitments made in the past. To paraphrase Merleau-Ponty, the fundamental, unified consciousness, classically identified as the cogito, is not an immaterial substance that assures identity; rather, identity and consciousness are accomplished in our lived-body experience of performing activities. By enacting our commitments in a personalized manner, we accomplish our identity. This last suggestion, I believe, highlights the mereological nature of personal identity, cross-methodologically as well as cross-culturally. From the rationalist approaches, both Western and Eastern, we find a strong sense and need for a universal, transcendental requirement, even if only a structural requirement, in order to make claims concerning identity, yet this will not provide for an account of a truly personal identity. From the empirical approach, we find contingency and particularity that demarcate a truly personal identity, while simultaneously encountering metaphysical difficulties in affirming such identity. The phenomenological account provided demonstrates the necessity for both universal and particular elements in determining personal identity, thus proving the interconnectivity of parts and wholes in affirming our personal identities.

## 9. Kant's Critical Concept of a Person: The Noumenal Sphere Grounding the Principle of Spirituality

*Claudia Bickmann*

### 1. Kant's Attempt to Unify the Extremes

Kant's critical concept of a person faced a double challenge: similar to the idea of Atman and Brahman within the Hindu Advaita-Vedanta philosophy, Kant presupposed the idea of a stable unity within us; but similar to Buddhist and Taoist critiques of this miraculous, "transcendent entity" that serves as a substantial ground within us, Kant shared the skepticism of Hume and Locke by refusing the idea of a stable substantial identity over time. Kant tried to reconcile the extremes. According to his Paralogisms of Pure Reason our "self" should no longer be regarded as a unifying metaphysical entity in Leibniz' sense; the only essences we might be aware of are the ever changing patterns of our historically and empirically bounded self—an ever-floating, never-resting appearance in space and time. Thus, as Kant claims, if we open the horizon for the unconditioned ground within us to understand the spiritual dimension of our personhood, we fail to understand the finitude of our empirical self; and if we stress the ever-changing character of our empirical self, no unconditioned ground within us may be found. Without access to the dimension of the absolute, as Kant claims, spirituality cannot be understood, and without regarding our empirical existence as bound to the spatiotemporal world, no concept of the objective world would be possible.

Kant's distinction between the "empirical and the transcendental I" opens the horizon of spirituality. The concept of the "transcendental unity of apperception" gives access to the spiritual dimension in its *formal* aspect. Since spirituality in a full sense, however, presupposes the unity of formal *and* material aspects, we will, in the second part of our analysis, open the horizon of Kant's principle of the "Ultimate

End of the Natural Dialectic of Human Reason" (A797/B825)<sup>1</sup>. First I will focus on Kant's concept of transcendence in its formal aspect. Secondly I will argue for Kant's concept of the all-integrating personality with regard to his "quasi-schema" of pure reason (A669/B697). Only with the "quasi-schematism" of pure reason (in the chapter Of the Ultimate End of the Natural Dialectic of Human Reason) does Kant succeed to reunite the material with formal idea of the absolute dimension of our personhood.

## 2. The *Formal* Constituent of Our Spiritual Self

### A. Synthesis presupposing "identity over time"

If knowledge acquirement presupposes identity over time, synthesis needs the unity of apperception to connect the empirical data according to the idea of an object as a "something in general = x". While receptivity provides only the *content* of our concepts of the empirical (or *a priori*) data in space and time, our understanding spontaneously grasps these manifold intuitions into one (act of) knowledge. Thus their identification as an object is due to a spontaneous act of our understanding rendering the *synthesis* of the manifold as necessary. This act of identification of an "object in general" then requires three fundamental acts of synthesis, whereby the manifold is "gone through, taken up and connected" (A99).

(a) While synthesis presupposes not only the manifold of intuitions, but also the ideas of connection and unity as the substratum of this unifying act, the fundamental act of the *apprehending synthesis* already gives the first hint of the predominance of an operating unifying *ground* within us. Kant explains this indispensable unifying function by counting larger numbers of elements. This operation "is a synthesis according to concepts" (A108) that can be executed only by reference to a common ground of unity—for instance, the decade. This concept (e.g., "decade") is what renders the synthesis of the manifold necessary.

The act that is immediately directed upon intuition is the act of synthesis of apprehension. Apprehending the manifold means "to order, connect und bring the empirical or pure elements into relation" (A99). Without connecting and bringing together the elements accord-

---

1 Translated by Norman Kemp Smith (London: Macmillan, 1929).

ing to a priori rules of our understanding, our empirical imagination would not find the opportunity to “exercise appropriately the affinity of the given object”. The act of name-giving already presupposes, as Kant claims, an “a priori rule”, whereby our empirical synthesis is possible, because if “one and the same thing named sometimes in one way, sometimes in another”, the reproduction of a certain name would be impossible (A100 f).

(b) Similar to the act of apprehending the manifold, there must be something that, “as the *a priori* ground of a necessary synthetic unity of appearances”, makes their *reproduction* possible” (A101). Thus, as Kant argues, “the synthesis of apprehension is ... inseparably bound up with the synthesis of reproduction.”

(c) Important for Kant’s assumption of the *a priori* unity of apperception, however, is the third synthesis, *recognition in a concept*. Recognition presupposes thinking of the same in different time sequences. Without this thinking of the same we would be unable to identify the present with the former state of our apprehension (A103). Only our consciousness can form the unity of a concept. If I forget the former unit while adding the latter, no single number and no total sum could be found. Thus, apprehending a total number makes the concept of the number necessary; in Kant’s view this is nothing else than the “consciousness of this unity of synthesis” (A103). Just by analyzing *the concept of a concept*, as Kant points out, we can conclude that such consciousness “must always be present”, even if it may often be only faint and not necessarily connected with the act itself. Without it, concepts (and so also, the knowledge of an object) would be impossible.

## B. The transcendental “I”

Speaking of a corresponding object can only be understood as referring to “something in general = X” (A104), since we have its representations by our synthetic actions. Since the relation of our knowledge to its objects cannot be arbitrary, it must follow such *a priori* rules, rendering the relation to a corresponding object as *necessary*. Thus our thoughts must not only be constituted by a unifying concept of an object, this unifying concept must also presuppose an internal principle of identity within us; the “transcendental I” as a unifying stable ground that enables us to synthesize different representations in a judgment, such as  $S = P$ .

Concerning this presupposed type of reference two moments are striking: (a) since we deal only with the representations of the given empirical data, it is evident for Kant that the respective object of our reference can only be the “unity of our consciousness in the synthesis of the manifold of representations” (A105); and (b) since the something we refer to cannot be distinct from our representations it can only be “the unity which the object makes necessary” (i. e., the “*formal* unity of our consciousness”). Kant finally concludes by stating: no object could be thought without a rule making the reproduction of the manifold of the empirical data *a priori* necessary. Those rules determining the connection of the manifold are named the *formal* object: the object in general. Thus objectivity, as the formal unity of our consciousness, is the result of our unifying actions that make the unity of apperception possible (A105).

The concept of an object, then, is not more than its representation = x by such *a priori* rules that are universal and serve to unify of the manifold of outer appearances. Now, if the concept of objectivity implies the subjective unity of apperception, as a unity over time that makes the different time sequences possible, we must presuppose a transcendental condition as its fundament (A106):

a transcendental ground of the unity of consciousness in the synthesis of the manifold of all our intuitions, and consequently also of the concepts of objects in general, and so of all objects of experience, a ground without which it would be impossible to think any object for our intuitions; for this object is no more than that something, the concept of which expresses such a necessity of synthesis.

### C. The person as synthesis of apperception and free will

Here we face the specificity of Kant's attempt to mediate the extremes; his theory of knowledge paves the way. Since necessity implies a transcendental condition as its fundament, the need to synthesize the empirical data renders necessary “a unity over time” as a transcendental unity. The numeric identity of the constituting self *and* the ever-changing empirical consciousness in a spatiotemporal world, as the two complementary dimensions of the same relation, are the reciprocal and irreducible aspects of our consciousness: dialectically intertwined like being and becoming, identity and change; two aspects of one act, as the unity of the extremes—the non-contradicting but complementing dimensions of



our consciousness. Thus, Kant's theory of knowledge calls attention to a unity in difference: both sides are intertwined as the stable identity connects and relates the ever-changing flux of the empirical representations.

Identity over time is not only an ingredient of our knowledge acquirement. As the unconditioned ground of our moral actions, it is also the preconditioned ground for our free will. Both, the identity within us in a theoretical sense and the unconditioned ground of our free actions, are united in a third dimension: the concept of a personality integrating the natural and supra-natural, the formal and the material dimensions in a non-contradicting way. The unconditioned stable ground within us gives rise to the noumenal sphere of transcendence, offering a hint to understand our spiritual self.

Kant's critical philosophy purifies this idea of the noumenal sphere: the transcendental meaning of the concept of an object, linked to the appearing phenomenon in space and time, is only "a representation, the transcendental object of which is unknown" (A109). But without any access to things in themselves, it should be similarly misleading to assume a unifying stable entity within us as a transcendent supra-natural entity.

This distinction counts for our theoretical and practical reason likewise: within the horizon of our theoretical consciousness we presuppose the unifying ground of our transcendental self; within the field of our practical reason we presuppose freedom as the transcendental principle of the categorical imperative. Both principles—the transcendental unity of apperception and the unconditioned ground of our free will—provide only the *formal* principles of our consciousness. The religious dimension of our consciousness, however, being bound to both dimensions of the absolute within us, presupposes a unifying formal *and* material principle. But how to find an unconditioned spiritual ground within us, if both the transcendental unity of apperception and freedom as the unconditioned principle of our practical reason are only formal principles, while spirituality is bound to the unity of both dimensions?

Concerning the highest principle of religion, the idea of a *divine being*, it is equally valid, that this idea of the all-integrating and all-sufficient ground of all beings cannot just be understood as a formal principle; it is likewise needed as formal and material "cause of all possible effects". But since all material causes are due to our empirical knowledge, they cannot be anticipated in any transcendental analysis. Thus the idea of a *divine being* as the final cause of our existence cannot be presupposed

as an in-itself-determined entity; it can only be *postulated*. Since it cannot be taken as a highest substance in itself, we may only postulate its existence as the principle of our pure reason “in the search for the unity of the grounds of explanation”. As Kant puts it (A696/B724 f): “We may indeed be allowed to *postulate* the existence of an all-sufficient being, as the cause of all possible effects, with a view to lightening the task of reason in its search for the unity of the grounds of explanation.”

### 3. Spirituality as the Integrative Dimension of Consciousness

#### A. How to have a coherent view of the concept of spirituality

Kant develops his idea of spirituality within the horizon of the idea of the “quasi-schematism” of pure reason. On a transcendental level the quasi-schematism of our pure reason reunites the material with the formal aspect of the absolute principle within us. While the identical stable unity within us—needed as the ground of the transcendental unity of apperception—paves the way to understand our spiritual self, the quasi-schematism of pure reason provides the principle uniting the formal and the material aspect of our spiritual personhood.

Let us now focus on Kant's idea of personhood involved in his concept of the “Ideal of the highest good” (A804/B832). I will argue that “the Ideal of the highest good”, regarded “as a determining ground of the ultimate end of pure reason”, opens the horizon of the integrative sphere between both the speculative and the practical employment of our reason. How is it possible to find a coherent concept of personhood: integrating sensibility, understanding, and reason in order to enable us to follow our natural and intelligible ends and finally to find access to the intelligible world?

Kant reflects this possibility of a moral and spiritual orientation on a transcendental level. Conceding the fact that pure reason is the capacity to act under moral rules, our access to a supra-natural world seems to be unquestioned. We are inhabitants of the spiritual world inasmuch as we find orientation in the sphere of nature. While our senses, as the grounding capacities of receptivity, give access to the natural world, our understanding—the capacity of connecting and relating the appearing data in space and time—provides the concept of an object in gen-

eral. Thirdly, our reason equips us with the capacity to open the horizon of the transcendental principles and ideas, even if we cannot infer by this capacity the existence of such in-themselves-contradictory entities like the unconditioned independent “substances over time” that integrate the ever-changing appearance of the empirical self or, alternatively, the all-determined world-integrating freedom of the will.

### B. Reason and systemization

Kant is skeptical with regard to our capacity to understand the area of transcendence. His idea is rather to unfold a coherent concept of our cognition that may harmonize with the claims of our practical reason. Therefore, he opens the distinction between different levels to approach the phenomenal and the intelligible world: While our senses are connected by the *understanding*, *reason* connects the empirical operations of our understanding to find its *systematic unity*. The systematization of our theoretical knowledge is the operational field of our reason.

Inasmuch as our understanding needs a schema to determine specific objects, the idea of our reason would be empty without an analogous determining schema of its principles. Thus our reason provides, in the same manner as our understanding, a *unity* “as regards the conditions under which, and the extent to which, the understanding ought to carry the systematic connection of its conceptions” (A664/B693). Not *the object itself* is addressed by this rule of the systematic unity of understanding, but the procedure to unite the operations of our understanding. *The only available object for our reason is the connecting activity of our understanding*. But according to Kant reason needs, similarly to the understanding, a quasi-schema that follows the principle “of a maximum of the division and the connection of our cognition in one principle”. It gives us the “rule or principle for the systematic unity of the exercise of our understanding” according to the idea of the unity of an object (A664/B693).

These maxims of reason, operating only regulatively on the exercises of our understanding, organize the two extreme tendencies of our reason: generalization and specification. The thing in itself as a completely determined object is now in Kant’s concept of pure reason the “*conception of a sum-total of reality*” (A664/B693), an *ens realissimum*, *determined by that predicate of all possible contradictory predicates, which indicates or belongs to being*.” Here we find Kant’s *fully determined concept of an individual being*; it

is the “*transcendental ideal of the pure reason*” (A576/B603), “*which determines the material condition of all existing beings. As the highest material condition of its possibilities it forms the preconditional ground to determine an object in a qualitative sense.*”

### C. “Of the Ultimate End of the Natural Dialectic of Human Reason” (A797/B825)

I will conclude by analyzing Kant's concept of personhood; this is not complete without a specific schema that may represent a fully determined object to our minds. Since the object obtained is nothing more than its specific relation to the idea of our pure reason, the idea of the completely determined object (i. e., “its objective reality”) does not consist in the fact that it has an immediate relation to an object; the idea of the all-determined object is merely a “heuristic schema” following the “necessary conditions of the unity of reason” (A664/B692 f). By this it indicates only the way to investigate the constitution of and the relations between the objects in our empirical world.

Only by reference to such a heuristic schema, the three transcendental ideas (psychological, cosmological, and theological) may produce systematic unity among our empirical laws. Kant's “quasi-deduction” of the three speculative ideas, restricted to their regulative function for the systematic unity of our empirical cognition, does not neglect the limits of our understanding—its restriction to the phenomena in space and time—by presupposing transcendent metaphysical entities, but it helps to complete the operations of pure reason.

(a) Here we find the Kantian solution to the problem of the above mentioned contradicting positions: either positing a stable unifying substance as the grounding principle of our selves or alternatively supposing that the only essences we might be aware of are the ever-changing patterns of our historically and empirically bounded self. As previously stated, Kant's schema of a person tries to harmonize the extremes: we may consider ourselves as an ever-floating, never-resting appearance in space and time; we are simultaneously permitted to assume that our empirical person is grounded in a stable unity over time. While our free will presupposes this unconditioned ground within us, our spiritual self mediates between the extremes of the conditioned sphere of our final empirical self and the unconditioned ground of a supra-natural self within us. But

how should this mediation succeed without neglecting the critical limits of our understanding?

Kant stresses the point that the unity of our person may possess only comparative reality and cannot be more than “the reality of a schema of the regulative principle of the systematic unity of all cognition” (A670/B698). Thus, the unity of a person is not regarded as an actual thing, but considered only in an analogous way to the idea of a completely determined object. The schema of a person can be understood as follows: we cogitate a person connected with all the contingent appearances, actions, and phenomena in space and time *as if* it would be a simple substance endowed with personal identity and possessing a permanent existence while all the states linked to its body continuously change in space and time.

(b) The same counts for the schema of cosmology: the heuristic schema of a cosmological analysis consists in an investigation of the conditions of the natural world and its phenomena, internal as well as external, “*as if* they belonged to a chain infinite and without any prime or supreme member,” but without denying “the existence of intelligible grounds of these phenomena” (A672/B700). Since this intelligible ground can never be an adequate object of our scrutiny, we just regard the appearing phenomena *as if* they might be grounded in a spontaneous, unconditioned intelligible principle.

(c) The same counts for the third transcendental idea: in the sphere of theology, as the all-embracing principle of the world as a whole and its grounding principle, we similarly have to regard the whole system of interconnected phenomena as a “dependent and sensuously-conditioned unity”; nevertheless, it simultaneously might be regarded as being based upon an all-sufficient supreme being as the ultimate ground, apart from the spatiotemporal world. Analogously to the principle of freedom as the constitutive basis for a moral world, we have to regard the supreme being as “self-subsistent” (A672/B700), equipped with “a primeval and creative reason, in relation to which we so employ our reason in the field of experience, *as if* all objects drew their origin from that archetype of all reason.”

#### 4. Conclusion

Thus the key points of Kant's quasi-deduction of the ideas of pure reason are the following: (a) we are not supposed to deduce "the internal phenomena of the mind from a simple thinking substance, but deduce them from each other under the guidance of the regulative idea of a simple being"; and (b) "we ought not to deduce the phenomena, order, and unity of the universe from a supreme intelligence, but merely draw from this idea of a supremely wise cause the rules which must guide reason in its connection of causes and effects" (A672/B700).

If we now have to consider the result of our analysis and try to figure out the constitutive principles of Kant's theory of a spiritual subject, we may state that, while Kant's concept of personhood gives access to the area of transcendence, it nevertheless affirms the ever-changing, never-resting character of our personhood in the appearing world. Thus, his concept of personhood integrates the polarity of our moral and our sensual world and, as a synthesis of the extremes, gives rise to our spiritual self: the spiritual self as characterized by its double nature in its sensual and supra-natural intelligible nature.

While Kant avoids extending our cognitions beyond the objects of possible experience, he nevertheless extends the empirical unity of our experience "by the aid of systematic unity" (A686/B714), viewed as schema of our pure reason (i. e., as a necessary regulative principle of our pure reason). Thus, transcendence may be regarded as an ingredient dimension of our pure reason and by this it is objectively valid.

Now we may conclude by indicating the consequences of our analysis: Kant avoids dichotomizing the extremes; both dimensions, the idea of a stable transcendent unity of our personhood and the idea of an ever-floating, never-resting personality in space and time, should rather belong to our personhood as the two opposite poles of our empirical and our transcendental self. He succeeds by clearly distinguishing between the different levels of our cognition: there is no Kant without the thing in itself; the thing in itself rather has to be maintained to safeguard the systematic ground for the idea of the absolute, for our moral and spiritual self, and—in the end—for the idea of the highest good as the in-itself-contradictory unity between our natural striving for happiness and our acting according to moral rules.

Thus religion is regarded as an inseparable dimension of our self. Similarly to the idea of a "transcendent entity" within Hindu Advaita-Vedanta philosophy, regarding the unconditioned self (Atman) as

the ultimate ground within us, Kant maintains the rationalistic view of a stable unity within us, and simultaneously affirms the Buddhist and Taoist critique of this miraculous “transcendent entity over time” by facing the arising skepticism of Hume and Locke, arguing that the only “essences” we might be aware of are to be found only within the flux of restless time-sequences.

## 10. Truth, Falsehood and Dialectical Illusion: Kant's Imagination

*Christine Lopes*

### 1. Intuitions as Non-Full-Fledged Conceptual Content

Kant claimed that the mind forms a veridical representation of spatio-temporal objects by bringing intuitions into a conceptual synthetic unity. He also claimed that the processes of synthesis and unification of representations are *a priori*, rule-governed by categories of thought. These two processes would thus constitute *a priori* conditions for objective knowledge. While Kant explained unification of representations as the function of a faculty of conceptualization, the “understanding”, he described the process of synthesis as a function of the faculty of imagination or *Einbildungskraft*. In what follows I consider in large brushes of argumentation that philosophers such as Kant, who are primarily concerned with propositional conditions of truth, must conceive of truth in relation to non-full-fledged conceptual content. One upshot of such a view is that valuing *imagining* and *error* must be a condition of being a rational person.

Kant famously argued in the *Critique of Pure Reason* for a double formal independence of the intuitions of space and time from the realm of empirical reasons (*viz.* from both sensory input and conceptual representation with perceptual content). This formal independence of spatio-temporal intuitions is properly expounded in the Axioms of Intuition, but already figures in the Transcendental Aesthetic, namely as the framework for Kant's method of enquiry into the role of intuitions in conceptualization.

In the transcendental aesthetic we will therefore first isolate sensibility by separating off everything that the understanding thinks through its concepts, so that nothing but empirical intuition remains. Second, we will then detach from the latter everything that belongs to sensation, so that nothing remains except pure intuition and the mere form of appearances, which is the only thing that sensibility can make available *a priori*. In this investigation it will be found that there are two pure forms of sensible in-



tuition as principles of *a priori* cognition, namely space and time, with the assessment of which we will now be concerned (A22).

My aim is to re-evaluate briefly the analytical entailment of judgment in general and the double formal independence of intuitions. My main goal will be to re-evaluate the function of intuitions in the structure of Kant's critical metaphysics.

One of the most daunting tasks for a Kant scholar is to account for this double formal independence of intuitions without crushing it under the analytical hammer. Even Robert Hanna's extraordinary interpretation of the Kantian *intuition* as non-conceptual mental content still does not manage to free it from analytical conceptualism.<sup>1</sup> This is evident from the conceptualist terminology that Hanna deploys in his analysis of Kant's concept of intuition: the latter is *indexical*, its objects are *veridical*, and so on. Intuition is a theoretical concept that Kant uses to explain the possibility of knowledge and is not analyzable in the same way as other concepts. Theoretical concepts are not appropriate to describe/explain/exemplify mental or physical events or objects. They belong in the space of reasons, as we say these days, not in the space of experience. For no other reason does McDowell in *Mind and World* step resolutely inside the Kantian realm of intuitions and declare them to be just as conceptual as any other mental content that can be used in justification of empirical judgments.

Hanna uses the conceptualist terminology to explicate the Kantian concept of intuition as all-versatile, virtually "blank-canvas" like elements of cognition: they are non-conceptual in four different hierarchical senses, according to Hanna, ranging from being a-conceptual, or lacking in concept entirely, to being conceptual without a self-conscious subject that thinks them. In this way Hanna obtains for intuitions the extraordinary cognitive property of "servants of any master", as it were, as they are described as non-conceptual matter for all modes of cognition—ranging, accordingly, from sensations to judgments. In a neat contrast, intuitions are with McDowell well-integrated conceptual components of communicable experience. I offer later on a brief account of McDowell's own interpretation of the Kantian notion of intuition, and my point of objection to his interpretation.

---

1 Robert Hanna, "Kant and Nonconceptual Content", *European Journal of Philosophy* 13.2 (August 2005), 247–90.

The question I wish to consider first is whether it is logically possible to describe any element of cognition as a non-full-fledged conceptual (*n-fc* for short) representation. Surely, the act of description would at once bring such an element under conceptual grasp. Or would it? Kant's own initial description does not really settle the matter (A20–1/B35):

If I separate from the representation of a body that which the understanding thinks about it, such as substance, force, divisibility, etc., as well as that which belongs to sensation, such as impenetrability, hardness, color, etc., something from this empirical intuition is still left for me, namely extension and form. These belong to the pure [*non-empirical*] intuition, which occurs *a priori*, even without an actual object of the senses or sensation, as a mere form of sensibility in the mind.

The passage describes a mentalist procedure for distinguishing sensation and conceptualization from intuitive representation of general spatio-temporal qualities. In what sense could the theoretical move of accounting separately for conditions of sensibility and conditions of conceptualization entail awareness of a *n-fc* factor in experience?

It is not difficult to imagine that what is initially presented in the Kantian text as a positive formal feature of cognition—namely, that we seem to be able to form *n-fc* representations (i. e., spatio-temporal representations of objects in the absence of sensory data or empirical conceptualization)—can also lead us into substantial error. For instance, take my quick judgment—based on my sensation of freezing hands as I unknowingly mistake boiling for cold water during the washing-up—that I have turned the cold water tap on. In this judgment I preserve the capacity to relate my sensation of freezing hands causally to the water that hits my hands, while mistaking what is only a physical/physiological property of my body—the sensation of freezing cold hands—for the actual physical state of the water that hits my hands. The preservation of the capacity in question can be seen as an instance of *n-fc* apprehension of spatio-temporal objects, made possible by what Kant would call empirical intuition: “empirical intuition [is intuition] of that which, through sensation, is immediately represented as real in space and time” (B146–7). My judgment that I have turned the cold water tap on contains an element of correctness, in that I correctly causally relate my physical sensation of freezing hands to the water that hits my hands. Reaching very low temperatures while in liquid state is certainly one of the physical properties of water. But my judgment is physi-

cally, albeit not psychologically, false. For it does not follow from my sensation of freezing cold hands that the water that hits my hands is cold.

Take now one's judgment—based on one's ignorance of non-Euclidean geometries—that the sum total of the internal angles of any triangle is  $180^\circ$ . In this judgment one preserves the general capacity to conceive a triangle while mistaking what is only a possible for a necessary mathematical property of triangles. The preservation of the capacity in question through mistake can be seen as an instance of *n-fc* apprehension of spatio-temporal objects made possible by what Kant would call mathematical intuition: “[in] mathematical ... indeed [in] geometrical construction ... I put together in a pure intuition, just as [I do] in an empirical one, the manifold that belongs to the schema of a triangle in general ...” (A718/B746). One's judgment retains an element of correctness to the extent that one attributes to triangles the property of having three internal angles; but one's judgment is mathematically false insofar as in non-Euclidean geometries the sum total of a triangle's internal angles is not  $180^\circ$ .

Now the two examples above *illustrate* a certain *n-fc* element in cognition, namely in a role of allowing for the correctness of our judgments even when the latter are based on a mistaken use of spatio-temporal representations. The question is how to *explain* such a *n-fc* element. Is it something we can become aware of through perceptual deception, illusion, or non-acquaintance with specific contents of knowledge? Or is it something that becomes intelligible to us in the first place because we already have knowledge of empirical processes of cognition? A classical illustration of the former view in philosophy is the belief that imagination is primarily involved in psychological deception and illusions—as in dreams, fantasies, and hallucinations. Kant did not hesitate to acknowledge the role played by imaginative processes in deception, illusion, and ultimately error.

But Kant also stepped out of the philosophical habit of associating imagination primarily with empirical deception, illusion, or lack of clarity in thinking. He claimed, in fact, that imagination is a mental process that crucially supports judgment. What can possibly be the rationale behind Kant's proposition of this dual role of imagination, its being involved in both error and correct judgment? I find in the first *Critique* a highly plausible answer to this question. The identification of deceptive or illusory inferential moves in our reasonings about objects, and, through this, the identification of an intermingling of *n-fc* elements in cognition, is a valuable means of self-criticism available to the

human rational mind, whereby it comes to realize the philosophical importance of justifying propositions, explicating concepts of objects, and confirming/adjusting/abandoning beliefs. Kant explicitly asserts the methodological value of deception, illusion, and error that seems to arise from intuitive and imaginative processes (A642–3/B670–1).

The relation between spatio-temporal intuition and imagination is as fundamental in Kant's enquiry into the possibility of *a priori* knowledge as it is misunderstood in its relevance to Kant's empirical realist stance. So a student often hears that the major difficulty with Kant's philosophy lies in his attempt to discover *a priori* rules for conceptualization of what is intuitive, and therefore, by definition, non-determinable by *a priori* principles. While this diagnostic cannot be seen as entirely inaccurate, it falls short of telling the whole story about Kant's critical philosophy. It is in my view quite precipitate to claim that Kant's analysis of a relation between intuitions and concepts must either be an intrinsically flawed analysis of the general concept of cognition into two opposing elements of cognition—*conceptual* and *non-conceptual* elements—or require no distinction at all, at least at the level of justification of propositions.

The discovery of *a priori* laws of conceptualization has to do in the first *Critique* with the discovery of truth conditions for empirical judgments, conditions that are not themselves empirical. Kant's philosophical point about knowledge of reality being tied down to an empirical use of *a priori* concepts has fundamentally to do with the practice of judgment, and, by implication, with the possibility of erroneous judgment. Kant's self-professed empirical realism seems to have to do, as a philosophical position, with understanding how we rationally overcome *error*.

## 2. Error and Imagination

Error belongs in judgment: it arises neither from perceptual misapprehension alone, nor from purely logical fallacies. As Kant puts it (A294/B350), “neither the understanding by itself (without the influence of another cause), nor the senses by themselves, can err.” To err involves an ability to conjecture about objects using principles of inference that are not as objective and valid as we think: “Error requires that we hold a false judgement to be true” (24:720). As error involves an ill-function of judgment that can be explained neither on the grounds alone of a mental distraction from perceptual evidence, nor on the

grounds alone of ignorance of logical principles of inference, Kant points out that the matter of errors is “neither true nor false” (24:721). Error may involve perceptual misapprehension, and it may involve logical fallacy, but error always arises from mistaken use of principles of inferences in the course of *judgment* (A293/B350):

Truth and illusion are not in the object, insofar as it is intuited, but in the judgement about it insofar as it is thought. Thus it is correctly said that the senses do not err; yet not because they always judge correctly, but because they do not judge at all. Hence truth, as much as error, and thus also illusion as leading to the latter, are to be found only in judgements, only in the relation of the object to our understanding.

Judgment is thus the sole *medium* of illusion and error, just as much as of truth. But what is the *source* of error? Some passages in the first *Critique* seem to suggest that Kant’s conception of error is either that of an inadequate influence of sensibility on cognition—through sensation, perception, or imagination—and/or a lack of attention to logical rules of inference. The following passages from the Dialectic seem to suggest just this.

Because we have no other sources of cognition besides [sensibility and understanding], it follows that error is effected only through the unnoticed influence of sensibility on understanding, through which it happens that the subjective grounds of the judgement join the objective ones, and make the latter deviate from their destination just as a moved body would of itself always stay in a straight line in the same direction, but starts off on a curved line if at the same time another force influences it in another direction. (A294–5/B350–1)

Logical illusion ... arises solely from a failure of attentiveness to the logical rule. Hence as soon as this attentiveness is focused on the case before us, logical illusion entirely disappears. (A296–7/B353)

But error has to do not only with logical ignorance or an inappropriate influence of sensibility, and of empirical imagination in particular (A295/B351–2), over *thinking*. Error has more fundamentally to do, as Kant hints it in the first of the passages above, with the act of mistaking for substantive mental acts (i. e., acts of cognition of real objects) what are merely formal acts of conjecturing and reasoning about the possibility of objects in general. This type of error interests Kant most from a philosophical point of view, as he sees it as being intrinsic to the speculative use of reason. He denounces it and deals with it in the Dialectic.

While inattentiveness to logical rules leads to logical illusion and can be avoided, the overestimation and misjudgment (in itself an act of judgment) of the objective scope of our judgments leads to transcendental illusion (A296–7/B353–4):

Transcendental illusion ... does not cease even though it is uncovered and its nullity is clearly seen into by transcendental criticism (e. g., the illusion in the proposition: “The world must have a beginning in time”). The cause of this is that in our reason (considered subjectively as a human faculty of cognition) there lie fundamental rules and maxims for its use, which look entirely like objective principles, and through them it comes about that the subjective necessity of a certain connection of our concepts on behalf of the understanding is taken for an objective necessity, the determination of things in themselves. [This is] an **illusion** that cannot be avoided at all, just as little as we can avoid it that the sea appears higher in the middle than at the shores, since we see the former through higher rays of light than the latter, or even better, just as little as the astronomer can prevent the rising moon from appearing larger to him, even when he is not deceived by this illusion.

Kant considers transcendental illusion a species of error—one that is unavoidable—thereby emphasizing its unsubstantial but nonetheless insidious nature. As some Kant scholars have already suggested, the notions of error and transcendental illusion are in a sense synonymous in the first *Critique*.<sup>2</sup> My contribution to this view can be put as follows: the way these two notions converge in meaning also defines a certain particular function for imagination in Kant’s critical philosophy. This is the function, as I argue next, of a psychological mediation between veridical and non-veridical representations, of a psychological guarantor of *possible* empirical reality for representations.

For Kant, imagination is not the source of error but is rather the only authoritative subjective source of mediation for conceptual ambitions that lack in empirical content fundamentally. The following passage illustrates this point well (24:710): “The more universal the understanding is in its rules, the more perfect it is, but if it wants to consider things *in concreto* then it absolutely cannot do without the imagination.”

Error, *qua* transcendental illusion, is with Kant an unredeemable (constitutive, as he calls it) inconvenience for reason, not for imagination. But then imagination must, by implication, somewhat benefit

---

2 Michelle Grier argued (in *Kant’s Doctrine of Transcendental Illusion* [Cambridge: Cambridge University Press, 2001]) for the crucial role of Kant’s doctrine of transcendental illusion in his critical-theoretical philosophy.

from error. Given Kant's claims concerning the function of imagination as an *a priori* act of synthesis of conceptual and *n- $\hat{c}$*  representations, we might be able to show that imagination somewhat benefits from instances of erroneous cognition (as in the washing-up and the non-Euclidean geometry examples); the latter reminds us of the *n- $\hat{c}$*  in cognition. Let us start with propositions that we can un-problematically infer from Kant's notion of transcendental illusion. We can infer that a critique of the non-empirical use of reason is a method whereby we become aware of cognitive ambitions that cannot be fulfilled by experience—hence *error*—but that are nevertheless at the heart of reason's job to set up, namely through its concepts (A310–1/B366–7):

The term “a concept of reason” ... already shows in a provisional way that such a concept will not let itself be limited to experience, because it deals with a cognition (perhaps the whole of possible experience or its empirical synthesis [of concepts and intuitions]) of which the empirical is only one part; no actual experience is fully sufficient for it, but every experience belongs to it. Concepts of reason ... deal with something under which all experience belongs, but that is never itself an object of experience; something to which reason leads through its inferences, and by which reason estimates and measures the degree of its empirical use, but that never constitutes a member of the empirical synthesis [of concepts and intuitions].

What are we to make of Kant's discovery of this form of illusion? Kant argues for its inevitability, but he also believes we can avoid the fallacies that stem from transcendental illusion. Is it that we become better thinkers if we bear in mind the constant pull of transcendental illusion that underlies our best efforts to rationalize inner life? How are we to avoid the metaphysical fallacies? We **cannot** avoid them by simply becoming aware of their fallacious logical structure, as Kant warns (A339/B397). One looks into the formal validity of inferences involved in transcendental illusion—inferences that confer, for instance, substantiality to the *I* as the subject of the purely formal act of thinking—and one can only learn that they stem from premises that stand in a non-resolvable dialectical relation.<sup>3</sup>

---

3 Let me briefly consider Kant's examination of the Cartesian proposition of self-consciousness. It says, to put it concisely and roughly, that without my thinking being no object can be known by me to exist as something that persists in time and space and thus no *a priori* predicate of substantiality can be ascribed by me to the concept of an object. I infer from this proposition of self-consciousness that I am an immortal soul (i. e., that I am a substance myself). Not, clearly, as a mere formal predicate of an object in general, but as the underlying ontological

### 3. Imagination and Rationality

Could imagination be for Kant the mental faculty that enables *a priori* ascription of an objectively valid reference of concepts to objects (appearances)? This question, eccentric as it may sound, is worth considering. The requirement of identifying and removing fallacies is one of rationality for our thoughts, and in a sense transcendental illusion is just as constitutive of demands of rationality as is the call to contain the theoretical excesses coming from such demands. However, we need Kant to be more specific about the mental power whose theoretical priority comes, in stark contrast to reason and sensibility, from managing the formal gap between *thinking* and *knowing*.

Imagination is for Kant this truth-related faculty of the mind, whose conception depends on the assumption that concepts and intuitions are formally heterogeneous. It is precisely because imagination is constituted of these heterogeneous forms of representing objects that (1) it is the power to represent illusory, non-veridical objects of either type, intuitive or ideal, just as much as veridical objects,<sup>4</sup> and (2) it attains either of these two representations of objects not by *intuiting* or *thinking/infering/judging*, but by bringing the two together and thereby doing it *by itself*. The theoretical priority of imagination consists in the uniqueness of its representational power; it consists in bringing forth the question of truth in a fundamental manner (A146–7/B185–7).

---

force, the *hypokeimenon* of anything that can ever exist for me. How can I prove my inference of a substantial soul from the proposition of self-consciousness? I cannot. I find myself trapped between two dialectical propositions. One proposition says that *I*, as the subject of thinking, am objectively (in self-consciousness) part of the act of thinking. The other proposition says that *I*, the subject of thinking, have an objective existence that can be examined and known apart from my own thoughts, namely as a substance. See A345–6/B403–4, B406–12.

- 4 My point is not, of course, that all acts of intuition and all acts of reason are acts of representation of illusory objects. My point is that intuition and reasoning are acts of the mind whereby illusory objects of knowledge are sometimes represented (as in the washing-up and the non-Euclidean geometry examples, in the case of intuition, or, as in the case of demands of reason, through the representations of an immortal soul, God, and a transcendental freedom to will to act in absolute independence of empirical conditions) and that objects can be so represented because of the very nature of the act in question, i. e., intuition or reason.



In light of my exposition of a relation between imagination and the critical demand for justification of concepts of reason, a number of new pressing questions arise. For instance, how does Kant account for imagination as this most dignified tool of rationality? How do we know when we are incurring metaphysical extravagance with our judgments? Which alarm bells, so to speak, must ring in our minds as we go through our inferential moves? What is it, in such cases, that in the cognitive grasp as it were *skids* and prevents us from having our experiences in accordance with rational principles?

Error happens where some form of “mismatch” between judgment and experience takes place, making it impossible for us to sustain a claim of knowledge. Admittedly, the description of this “mismatch” is not an easy task in the context of transcendental philosophy. As Sebastian Gardner points out, transcendental illusion is not brought forth by the same inferential processes that bring forth “empirical illusion, the results of sensory deception, and logical illusion (due to inattentive application of rules of inference).”<sup>5</sup> What I now want to suggest is that we might not mistake Kant if we regard these other kinds of illusions as forms of illusion that, while not being brought forth by the same processes that bring forth transcendental illusion, do share the same final result with the latter. We would not necessarily misinterpret Kant’s critical stance by agreeing that error arises only where the realm of reasons fails to provide justification for empirical judgments.<sup>6</sup> In fact, this description of error seems to me perfectly applicable to transcendental illusion, just as much as to the other forms of illusion mentioned above.

What is error in the context of a failure or impossibility to provide justification for empirical judgments? Error concerns, in this context, something that in our empirical cognitions escapes through the conceptual net we deploy in our efforts to *judge*. Why should Kant be interested in devising *a priori*, necessary and universal conditions of possibility of experience, if not because he could see that error—rather than the ca-

---

5 Sebastian Gardner, *Kant and the “Critique of Pure Reason”* (London: Routledge, 1999), 215.

6 Examples of such a rational failure abound, and constitute the distinct subject-matters especially of ethics (justification of the legality or applicability of laws on ideas of freedom, civility, etc., rather than on arguments concerning human nature and custom), philosophy of religion (justification of the concept of a world created by an almighty and infinitely good Being and yet full of evil, human as well as natural—the so-called problem of evil), and philosophy of science (observational vs. theoretical claims).

capacity to determine truth or falsehood, which may or may not require experience—is constitutive of judgment? He makes this point clear in an ironic sociological remark (24:720): “There are more errors in an academy of sciences than in a village full of farmers, because more judging occurs there—he who never judges will never err either.”

#### 4. Final Remarks

Apart from the trivial fact that empirical reality seems to involve an element of contingency, of the unexpected as much as of the unknown, and thereby some degree of error and illusion in its cognition—what reasons do I have to believe that Kant explains the possibility of empirical reality for representations through an appeal to the notion of *in-c* elements in cognition? The answer to this question can at least in part be obtained by contrasting my interpretation of the nature and role of intuition in Kant’s first *Critique* with the interpretation given by John McDowell in *Mind and World*. My basic belief is that McDowell’s reading of Kant is problematic in that it interprets Kant’s philosophical intentions without taking into account the philosophical context of his allegiance to empirical realism.

Following Wilfrid Sellars’ interpretation of Kant’s notion of intuition in *Empiricism and the Philosophy of Mind*,<sup>7</sup> McDowell claimed that the Kantian spatio-temporal intuition is a species of conceptualization. This interpretation of intuition is necessitated by McDowell’s own attempt to examine and dismantle what Sellars first called the Myth of the Given. The Myth of the Given is the ungrounded philosophical belief in a raw, non-conceptual influence of the external world upon our cognitive apparatus. The Myth of the Given is the response, so McDowell argues, to the very specific methodological fear that, unless we postulate the status of the sensibly “given” for mind-independent objects, we may not be able to justify the idea that we are objectively free in the rational act of judgment.

As the Myth goes—or so McDowell denounces—we seem to think that there can be no objective experience of mind-independent objects, and consequently, no rational justification of empirical judgment, unless

---

7 Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Cambridge, Ma.: Harvard University Press, 1997); originally published in *Minnesota Studies in the Philosophy of Science* 1 (1956).

we postulate a “given” status to the worldly things. He urges his reader to realize once and for all that

The relevant conceptual capacities are drawn on in receptivity ... It is not that they are exercised on an extra-conceptual deliverance of receptivity. We should understand what Kant calls “intuition”—experiential intake—not as a bare getting of an extra-conceptual Given, but as a kind of occurrence or state that already has conceptual content. In experience one takes in, for instance sees, that things are thus and so. That is the sort of thing one can also, for instance, judge.<sup>8</sup>

Why should one have reservations about the Sellars–McDowell claim? It seems clear from my earlier considerations on the presence of a *n- $\hat{c}$*  element in cognition that there is neither formal nor substantial conflict between Kant’s account of intuitions as *n- $\hat{c}$*  representations and the claim that justification of cognition in general, including empirical *judging*, involves our awareness of instances in the process of conceptualization that are occasionally devoid of factual content.

I believe it is in proper philosophical interest to acknowledge that theoretical justification of claims of knowledge about mind-independent objects must involve the notion of what I have called a *non-full-fledged conceptual* element in cognition. We wake up to this element as traditional, scientifically proved, or merely habitual ways of understanding phenomena and empirical events *err*.

Finally, I hope to have offered a moderately convincing theoretical viewpoint for appreciating Kant’s theory of imagination. The idea that imagination brings the *n- $\hat{c}$*  or intuitive under a *a priori* rules of conceptualization belongs in a larger philosophical and methodological picture. It is not a mere coincidence that, out of all mental processes, the most subjective of them all—*imagining*—should serve the justificatory demands of reason in both Kant’s epistemology and criticism of rational metaphysics.

---

<sup>8</sup> John McDowell, *Mind and World* (Cambridge, Ma.: Harvard University Press, 1996), 9.

## 11. Persons as Causes in Kant

*Wolfgang Ertl*

### 1. Acting in Space and Time

Kant claims that human persons, by virtue of their rationality, can and should make a causal difference in the world of appearances. As he emphasizes particularly throughout the *Kritik der praktischen Vernunft*, they should shape the world of appearances and thus ultimately contribute, as much as they can, to making the highest derivative good a reality in the world. Even if, in Kant's opinion, this cannot be done without assuming the existence of God and his co-operation or *concursum*, it is certainly not something that could be done by God himself. Human beings must work themselves up to the state of virtue and, on pain of undercutting their freedom, nobody else can do that for them. In other words, cooperation from both types of agents is required for Kant, and therefore the realization of the highest derivative good must be something like a joint venture; this in turn is nothing other than an expression of Kant's metaphysical equality thesis in ethics.

But how is this possible given that he also holds that the world of appearances is causally closed and that reason is not natural in the sense of the Transcendental Analytic? According to Kant, by virtue of rationality and intelligence we are members of a world different from the one of appearances. Even if we concede that the second analogy might not by itself generate the full set of special causal laws valid in the world of appearances, the doctrine of the regulative use of reason makes it clear that we must not resort to causes outside of the realm of experience in our explanation of appearances.

Causal closure and the non-naturalness of reason seem to leave only two options: causal inertness of reason or systematic overdetermination. Causal inertness of reason is what a reading of Kant as a Davidsonian anomalous monist *avant la lettre*, as suggested by Hud Hudson<sup>1</sup> for ex-

---

<sup>1</sup> Hud Hudson, *Kant's Compatibilism* (Ithaca and London: Cornell University

ample, amounts to. For all its apparent similarities to a two-aspect (“TA”) reading of transcendental idealism (“TI”), anomalous monism does portray mental events (insofar as they are mental events) not as standing in causal relations to physical events, but as token-identical to physical events. The mental properties themselves play no genuine causal role.

Causal overdetermination is usually seen as a faulty notion; that is to say, most commentators claim there is no (or even cannot be any) genuine causal overdetermination. There is a broad consensus about this in particular in the philosophy of mind; if a position implies overdetermination, this is regarded as a ground for rejecting it. In this vein, recent attempts to read Kant’s claims along the lines of overdetermination take overdetermination not to be genuinely causal, but interpret overdetermination of, say, an event or action in terms of causes on the one hand and reasons on the other. Christine Korsgaard and Ermanno Bencivenga have tried this recently, with Korsgaard either making no use at all of TI or interpreting it in the familiar anodyne way, whereas Bencivenga draws on a rather strong reading of TI, based on the doctrine of object constitution, whereby the respective framework of regularities is what turns things into what they are.<sup>2</sup> An action can be constituted on the one hand by the frame of natural causality (spatiotemporal regularities) and within the framework of reasons (the regularities of rational beings) on the other hand. Thus, an action being overdetermined means it is causally determined in one framework and its transworldly identical counterpart is accountable in terms of reasons or rational regularities. Evidently, for all its merits regarding the problem of causal overdetermination, this strategy does not help us come to terms with how something non-natural can make a causal difference *within* nature. Tellingly, Korsgaard reads Kant’s take on the freedom and determinism issue along the lines of Peter Strawson’s so-called “new compatibilism”, whereby

---

Press, 1994). See Wolfgang Ertl, “Hud Hudson: Kant’s Compatibilism”, *Kant-Studien* 90 (1999), 371–84 for a more extensive discussion of this strategy.

2 Christine M. Korsgaard, “Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations”, in Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 188–221; and Ermanno Bencivenga, *Ethics Vindicated: Kant’s Transcendental Legitimation of Moral Discourse* (Oxford, New York: Oxford University Press, 2007).

both perspectives on action are indispensable even if they might be incompatible on the level of theory.<sup>3</sup>

In this essay I shall address the problem of rationality making a causal difference in the world of appearances by taking a look at Kant's discussion of the freedom and determinism problem in the *Kritik der praktischen Vernunft*. I shall proceed as follows: First I will sketch the structure of Kant's argument and elucidate where in the proposed solution the distinction of things in themselves and appearances comes in. Then, by taking a look at recent debates concerning a metaphysical or ontological TA reading of TI, I shall try to argue that a promising way of understanding Kant is to take him as regarding natural causal powers of agents<sup>4</sup> in a way similar to—but at least in one important respect different from—Leibniz' doctrine of *phenomena bene fundata*. To be sure, reason making a causal difference in the world of appearances is not sufficient for solving the problem of freedom, but it is an important part of such a solution. Conversely, since reason making a causal difference in this world is part of the solution to the problem of freedom, it is fitting to deal with this issue in the context of freedom in the first place.

## 2. Transfer of Powerlessness?

The section “Critical elucidation of the analytic of pure practical reason” (5:89–106) is unique in Kant's writings and one of the key passages in the *Kritik der praktischen Vernunft*, especially as far as the multifaceted relationship between theoretical and practical philosophy is concerned. It consists of 19 densely argued paragraphs and can be divided into three parts. Part 1, from paragraph 1 to 7, recapitulates the differ-

---

3 Peter Strawson, “Freedom and Resentment”, in Peter F. Strawson, *Freedom and Resentment and Other Essays* (London: Routledge, 1974), 1–25.

4 Without further argument, I am following Eric Watkins, *Kant and the Metaphysics of Causality* (Cambridge: Cambridge University Press, 2005) in his claim that according to Kant's metaphysics of causality, it is substances endowed with causal capacities which function as causes, and not events. This reading has, of course, far reaching consequences because it disconnects Kant from most of the contemporary theories of causality and associates him more closely with the Aristotelian tradition. Watkins uses the term “power” instead of “capacity”, partly because he wishes to put the emphasis on individuals, i. e., individual substances. Although I agree that it is the individual substance which functions as a cause, I nonetheless use the term “capacity” here in order to distinguish it from the notion “to be in one's power”, as we shall see below.

ences regarding the priority of concepts and principles and the parallel regarding the facticity of pure sciences and the fact of reason. Part 2 (paragraphs 8–17) deals with the problem of freedom and determinism. Part 3 (paragraphs 18–19) explores the fruitfulness of the concept of freedom in supplying surrogates for theoretical knowledge, in response to what for Kant are key questions of metaphysics.

The second part (i. e., the freedom and determinism section) is itself divided into two subsections. The first one (paragraphs 8–12) deals with the problem of freedom and causal determinism or natural necessity, the second (paragraphs 13–17) is concerned with the problem of human freedom and the creation of the world (i. e., its substances, in so far as they are actual) through God. Regarding both problems, Kant claims that TI is indispensable for providing a solution. As we shall see, with respect to the issue of natural necessity, TI is supposed to open up a possibility that natural necessity is not applicable to certain features or determinations of an entity. Concerning the problem of creation, the transcendental idealist doctrine of space and time as forms of human intuition undercuts the need for, or the implications of, regarding God as the ultimate causal source of our actions.

In what follows, I shall focus on Kant's proposed solution to the problem of physical determinism; but before turning to the details, it is worth mentioning that considerations regarding intellectual intuition (i. e., the intuition of a divine intellect) are employed (at least hypothetically) to render the transcendental idealist strategy intelligible (see 5:99). In this regard at least, the divine intellect is not considered as a threat to freedom, but as a device to secure it. (This is also connected to Kant's theory of conscience, but I shall leave out this feature of Kant's position here.)

Kant states the problem of freedom and natural necessity or causal determinism as follows, starting with the position of transcendental realism ("TR"):

Nimmt man nun die Bestimmungen der Existenz der Dinge in der Zeit für Bestimmungen der Dinge an sich selbst (welches die gewöhnlichste Vorstellungsort ist), so läßt sich die Notwendigkeit im Kausalverhältnisse mit der Freiheit auf keinerlei Weise vereinigen; sondern sie sind einander kontradiktorisch entgegengesetzt. Denn aus der ersteren folgt, daß eine jede Begebenheit, folglich auch jede Handlung, die in einem Zeitpunkte vorgeht, unter der Bedingung dessen, was in der vorhergehenden Zeit war, notwendig sei. Da nun die vergangene Zeit nicht mehr in meiner Gewalt ist, so muß jede Handlung, die ich ausübe, durch bestimmende Gründe, *die nicht in meiner Gewalt sind*, notwendig sein, d. i. ich bin in dem Zeitpunkte, darin ich

handle, niemals frei. Ja, wenn ich gleich mein ganzes Dasein als unabhängig von irgend einer fremden Ursache (etwa von Gott) annähme, sodaß die Bestimmungsgründe meiner Kausalität, sogar meiner ganzen Existenz gar nicht außer mir wären, so würde dieses jene Naturnotwendigkeit doch nicht im mindesten in Freiheit verwandeln. Denn in jedem Zeitpunkte stehe ich doch immer unter der Notwendigkeit, durch das zum Handeln bestimmt zu werden, was nicht in meiner Gewalt ist, und die *a parte priori* unendliche Reihe der Begebenheiten, die ich immer nur nach einer schon vorherbestimmten Ordnung fortsetzen, nirgend von selbst anfangen würde, wäre eine stetige Naturkette, meine Kausalität also niemals Freiheit.<sup>5</sup>

We can formalize this argument in a rather simplified manner as follows. There are basically two readings; I shall call them “A” and “B”, respectively:

*Reading A:*

1.  $PLp(t_1)$
2.  $N(p(t_1) \rightarrow p(t_2))$
3. Conclusion:  $PLp(t_2)$ ; 1, 2 and TPL under natural necessity

*Reading B:*

1.  $N(p(t_1))$
2.  $N(p(t_1) \rightarrow p(t_2))$
3. Conclusion:  $Np(t_2)$ ; 1, 2 and TNP
- [4.  $Np(t_2) \rightarrow PLp(t_2)$ ]
- [5.  $PLp(t_2)$ ; 3, 4 and MP]

---

5 5:94 f (following the Vorländer text): “Now, if one takes the determinations of the existence of things in time for determinations of things in themselves (which is the most usual way of representing them), then the necessity in the causal relation can in no way be united with freedom; instead they are opposed to each other as contradictory. For, from the first it follows that every event, and consequently every action that takes place at a point of time, is necessary under the condition of what was in the preceding time. Now, since time past is no longer within my [power], every action that I perform must be necessary by determining grounds *that are not within my [power]*, that is, I am never free at the point of time in which I act. Indeed, even if I assume that my whole existence is independent from any alien cause (such as God), so that the determining grounds of my causality and even of my whole existence are not outside me, this would not in the least transform that natural necessity into freedom. For, at every point of time I still stand under the necessity of being determined to action by *that which is not within my [power]*, and the series of events infinite *a parte priori* which I can only continue in accordance with a predetermined order would never begin of itself: it would be a continuous natural chain, and therefore my causality would never be freedom.” I have followed Gregor’s translation, but rendered the German “Gewalt” as “power” instead of her “control”, since “Gewalt” has been the established translation for “potestas” and “potentia” in Baumgarten’s *Metaphysica* (see §832 and §708, respectively), Kant’s text of reference in matters of ontology.



These formal accounts can be paraphrased as follows. On both readings the argument basically has the form of a hypothetical syllogism, with the same 2nd premise, namely a hypothetical proposition that is naturally necessary and connects two propositions that describe the state of the world at different times, respectively: the antecedent proposition a state of the world in the past, the consequent proposition a state of the world in the present or future. The two readings differ, however, with regard to the first premise. According to reading A, the agent is powerless with regard to facts of the past, whereas according to reading B, past facts are naturally necessary. On reading A, premise 2 transfers powerlessness (“TPL”); on reading B, premise 2 transfers natural necessity via the transfer of necessity principle (“TNP”). Since for Kant, natural necessity implies powerlessness, powerlessness with regard to something present or future in reading B is derived by means of a fourth and fifth step.

With regard to “powerless” or “not in one’s power”, it is striking that Kant operates with an essentially negative notion here, while he closely associates, if not identifies “freedom” with being “in one’s power”, so that we need to read him *ex negativo*, as it were. He has two points in mind: (i) the one who has something in their power (i. e., in our case, the rational agent) must have causal capacities with regard to the thing that is supposed to be in their power; (ii) the rational agent must have options for exercising the causal capacity, that is to say, agents must at least have the option not to exercise their causal capacity. Moreover, Kant clearly thinks we are not in power with regard to the past, but he does not elaborate why he thinks so. Arguably, this is because we have no causal capacity extending backwards in time, and we have no options regarding past facts, because facts of the past are, in a sense, necessary. Finally, the problem with both transfer principles seems to be that they take away our options about the future, although we retain, of course, the causal capacity to influence the future. With regard to the transfer principles, something we are powerless over triggers or activates our causal capacities. In other words, the transfer principles do not imply occasionalism with regard to the future, whereby we are causally inert in all respects.

In the paragraphs following the passage under consideration here, Kant makes it clear how he does not wish to solve the problem either explicitly or implicitly. He does not wish to undercut the validity of premise 2 as far as all temporal entities or temporal properties of an entity are concerned, nor does he wish to deny the validity of the crucial

principle that powerlessness is closed under natural necessitation (from reading A); or from reading B, that TNP is valid. As we shall see, though, he is eager to uphold the validity of the transfer principle in one interpretation of “powerlessness” only, and he tries to carve out a second conception of both powerlessness, and in one’s power; for this conception, the respective transfer principle is not valid, nor does the implication from natural necessity hold.

Kant then briefly discusses the example of theft, taken up again later, in a way that may at first view look as if he endorses Peter Strawson’s version of new compatibilism. The upshot of Kant’s example is that, determinism notwithstanding and in full view of the thief’s powerlessness (in one sense of “powerlessness”, as it will turn out) to refrain from stealing, we hold the person doing the stealing responsible. In contrast to Strawson, however, Kant does not merely state that both takes on human behaviour are indispensable, even if they are at odds with each other on the level of theory. Rather, for him, there must be a way of reconciliation on the very level of theory itself.

What is equally important, at least for now, is that Kant also opposes the strategy of linking freedom with the causal source of action being internal, evidently having Leibniz in mind. For Kant, it does not matter whether the transfer of powerlessness or necessity occurs through internal or through external necessitating causes. Nonetheless, he may be a bit too quick to dismiss the internal origin requirement, since in his account it is most likely a necessary condition for freedom. At any rate, this is his position in *Grundlegung* III.

Moreover, Kant also makes it clear that in TI we are still powerless with regard to the past, that is to say, he does not claim that the transcendental idealist theory of space and time undercuts in any sense the traditionally so-called accidental necessity of the past. Hence, the TI theory of time is still an A-theory of time, according to which real change occurs. In other words, TI neither renders the past wholly non-necessary, nor supports what one could call Ockhamist strategies, whereby there are “soft facts” about the past (i.e., facts that depend on something not belonging to the past).

Finally, in TI natural necessity and freedom remain “widerwärtige Begriffe” (“mutually repellent concepts”, in Gregor’s translation); that is to say, an agent cannot be free and his action naturally necessary in the same respect. Again, however, and similar to the different conceptions of powerlessness, this will turn out to be true only with respect to

one conception of freedom (namely, as a power with options within nature, given the special laws as they are).<sup>6</sup>

So far we have just seen how Kant does think a solution cannot be found. The constructive side of his approach is laid out in two steps, and in his opinion, as mentioned above, TI can guide us out of this labyrinth. These two steps are: (a) natural necessity is not all-encompassing, there being a range of features of agents to which it does not apply; and (b) temporal determinations are the effect (“Folge”) of the agent’s intelligible causality.

### 3. Things in Themselves, Appearances, and Powers

Both steps in Kant’s constructive solution to the freedom–determinism problem take us deep into the core of TI and moreover into the infamous distinction between things in themselves and appearances. Here Kant obviously seems to endorse a TA reading in the context of rational agency. There are basically two variants of a TA reading of TI: an epistemological and a metaphysical, or ontological, variant. Whichever we choose, it is clear that Kant conceives things in themselves and appearances as forming what Ralph Walker has called, “composite wholes”.<sup>7</sup> Crucially, this distinction does not apply to actions, but to the rational agent whose action is under consideration.

Moreover, the distinction between the thing as it is in itself and the thing as appearance is crucial for applying the “widerwärtigen Begriffe” of freedom and natural necessity to one and the same entity. Kant claims:

... sondern jede Handlung und überhaupt jede dem inneren Sinne gemäß wechselnde Bestimmung seines Daseins, selbst die ganze Reihenfolge seiner Existenz als Sinnenwesens, ist im Bewußtsein seiner intelligibelen Existenz nichts als Folge, niemals aber als Bestimmungsgrund seiner Kausalität als *Noumens*, anzusehen.<sup>8</sup>

6 Admittedly, this is already an interpretation and some passages are consistent with such a Davidsonian strategy.

7 Ralph C. S. Walker, “The Number of Worlds in Kant”, (updated August 2007) <<http://www.philosophy.ox.ac.uk/members/rwalker/index.htm>>, accessed 24 January 2008.

8 5:97 f. „... but every action – and in general every determination of his existence changing conformably with inner sense, even the whole sequence of his existence as a sensible being – is to be regarded in the consciousness of his intelligible

In order to become clear about this relationship of determining ground and consequence, I will turn to a recent discussion, focusing on a metaphysical reading of the “composite whole approach”. This is the debate between Rae Langton and Lucy Allais.<sup>9</sup> Langton construes Kant’s distinction between things in themselves and appearances in terms of extrinsic and intrinsic properties of things. To facilitate matters, I shall speak of the intrinsic properties and extrinsic properties of one thing here, although it may not be easy to isolate such an individual in each case. Moreover, Langton claims that Kant subscribes to an anti-Leibnizian position in that he takes the extrinsic properties to be irreducible to the intrinsic properties (i. e., for her, extrinsic properties, such as relations, do not supervene on intrinsic properties, as they evidently do for Leibniz). By virtue of this anti-Leibnizian claim, the phenomena do not offer us a window into the world of things in themselves. In other words, Kant—on Langton’s reading—rejects Leibniz’ doctrine of the *phenomena bene fundata*. In this way, receptivity leads to humility, as she calls it (i. e., our inability to have substantial knowledge of things in themselves).

The criteria Langton uses for distinguishing extrinsic and intrinsic properties are twofold: (a) the so-called loneliness criterion (i. e., properties are intrinsic if they can be the only entities existing); (b) the criterion that these properties remain the same under variation of natural laws, if they are to be intrinsic. Criterion (b) has the important consequence that from her reading, intrinsic properties are inert as to natural causality, assuming that natural laws govern instances of natural causality or even describe causal powers.

The same criterion (b) is one of the foci in Allais’ criticism of Langton’s approach. She calls Langton’s inertia thesis regarding intrinsic, in particular essential properties (since evidently essential properties are intrinsic), her “modal intuition”. For Allais, this modal intuition is on the one hand a logical possibility, but on the other it is (i) at odds with Kantian doctrines and (ii) implausible on systematic grounds relating to the philosophy of perception. In short, for her, intrinsic properties are not

---

existence as nothing but the consequence and never as the determining ground of his causality as a *noumenon*.”

9 Rae Langton, *Kantian Humility: Our Ignorance of Things in Themselves* (Oxford: Clarendon Press, 1998); and “Kant’s Phenomena: Extrinsic or Relational Properties? A Reply to Allais”, *Philosophy and Phenomenological Research*, LXXIII (1) (2006), 170–85; and Lucy Allais, “Intrinsic Natures: A Critique of Langton on Kant”, *Philosophy and Phenomenological Research*, LXXIII (1) (2006), 143–69.

inert; rather what Kant was aiming at with his distinction between things in themselves and appearances is similar to different ways of referring to properties, namely as an opaque versus a transparent quality.

As far as the clash with Kantian doctrines is concerned, Allais claims that Langton's strategy renders special causal laws (as opposed to the causal principle of the second analogy) wholly contingent (though this might be just an unrelated second feature of her account), whereas Allais takes them to be necessary in the Kantian framework. The question is, of course, in what sense of "necessary", but we will come to this shortly.

With regard to Allais' systematic considerations, it is at least very doubtful whether considerations of the philosophy of perception can really help us understand the causal inertness of essential properties in the relevant sense. Here Kantian considerations interfere, since Kant distinguishes at least two types of the "nature" of a thing:<sup>10</sup> nature in the sense of the Transcendental Analytic, as only "comparatively inner" (cf. A277/B333) and concerned with theoretical entities, and the absolute nature of things, whereof we can have no knowledge, except that there must be such a feature of things. If considerations of perception yield anything at all helpful, then it is as far as nature in the sense of the Transcendental Analytic is concerned. Put differently, the distinction between different ways of referring to properties concerns the world of appearances itself.

Let us leave that point there and turn to the other issue. As we have seen, Allais claims that, in an important sense, the special laws of nature are necessary and it was left open above, for the time being, which sense of "necessary" is relevant here. Evidently, it cannot be "logical necessity", nor "metaphysical necessity", for in these senses every conceptually and really possible world must contain the same set of laws. Following Robert Hanna,<sup>11</sup> she claims that the special causal laws of nature are necessary in the sense that all worlds that contain matter have the same set of special causal laws. This is one species of the "necessity" related to synthetic *a priori* propositions or in short "synthetic necessity", the other species being the necessity pertaining to mathematical judgements.

---

10 Cf. the very important hint in *GMM* III (4:447), where Kant speaks of the "Natur der Sinnenwelt"; he seems to contrast this with a different nature, with rational nature being an example.

11 Robert Hanna, *Kant and the Foundations of Analytic Philosophy* (Oxford: Clarendon Press, 2001), 260.

But this again seems too strong: it is true that in every world that contains matter *the law of gravity* must hold, but there is certainly no merely deductive relationship between this, as it were, most general of the special causal laws, and all the other, more specific or fine-grained special causal laws. Rather, as Michael Friedman has suggested,<sup>12</sup> there is an element of (synthetic or real) contingency in the special causal laws detectable by experience; what is required is to ground these empirically discoverable regularities in more general laws, and there are a number of different ways these regularities might look, in particular the more specific they become in the hierarchy of laws. In short, the fundamental physical laws can differentiate in a number of ways. But if this is so, and since natural necessity plainly hinges on the actual set of special causal laws, logical space is opening up for a conception of power that cuts across different sets of special causal laws.

With this in mind, we can come back to Langton's criteria for distinguishing extrinsic and intrinsic properties. The loneliness criterion seems unproblematic, and by means of the variability of the special laws of nature, we have opened up the possibility to allow intrinsic properties, in principle, to remain the same while at least certain strata of the laws of nature are, or can be, changing. (Possibly, though, these latter considerations also apply to comparative inner natures.) However, moving in the opposite direction to that of Langton, our considerations so far also open up a promising route as to how an account of reason making an independent difference in the world of appearances can be construed. To see this, let us turn to an important line of criticism developed by those who are in principle quite sympathetic to ontological or metaphysical approaches, such as Ameriks and Falkenstein.<sup>13</sup>

By endorsing a global irreducibility claim, these critics say Langton has undercut the vital link between reason and the world of appearances, and along with it Kant's account of rational agency. This is indeed a valid point; it is surprising that she has not considered the issue of free-

---

12 Michael Friedman, *Kant and the Exact Sciences* (Cambridge, Ma., and London: Harvard University Press, 1992); and "Metaphysical Foundations of Newtonian Science", in Michael Friedman, *Kant and the Exact Sciences* (Cambridge, Ma., and London: Harvard University Press, 1992), 136–64.

13 Karl Ameriks, "Kant and Short Arguments to Humility" in Karl Ameriks, *Interpreting Kant's Critiques* (Oxford, New York: Oxford University Press, 2006), 135–57; and Lorne Falkenstein, "Critique of Kantian Humility", *Kantian Review* 5 (2001), 49–64.

dom in her book. But it is at the same time easy to see that this problem is repairable and, moreover, that this repair is also crucial for our question at issue.

In short, the most promising line to account for reason making an independent difference in the world of appearances is construing reason's causality in terms of supervenience. In the case of rational beings, the assumption of freedom requires that their natural causal powers supervene on their intrinsic properties in the absolute sense of the term "intrinsic". Clearly, rationality is one of their essential properties. There is no theoretical proof, however, for this supervenience; all Kant is required to say is that, in so far as we assume freedom, we need to buy the supervenience account, and that transcendental idealism provides the framework for rendering this possible.<sup>14</sup>

Of course, we need to be clear about the exact type of supervenience itself, the supervenience base and the supervening properties. For lack of space I will postpone this to a different occasion and discuss instead an important objection to this reading. According to this objection, supervenience talk does not license the assumption of a second form of causality such as causality of freedom or transcendental freedom,

---

14 Although I am not concerned with Kant's full argument for his version of compatibilism, a few words are in order concerning Watkins's recent suggestion in chapter 5 of his book (*Kant and the Metaphysics of Causality*, 301–61) mentioned above. I agree that Kant pursues a strategy one could call *altered law compatibilism* (as I suggested in "Schöpfung und Freiheit. Ein kosmologischer Schlüssel zu Kants Kompatibilismus", in Norbert Fischer (ed.), *Kants Metaphysik und Religionsphilosophie* (Hamburg: Meiner, 2004), 43–76). By this means Kant can undercut van Inwagen's consequence argument, in that laws of nature (special causal laws of nature, that is) are a function of our free actions or at least depend on our free actions (cf. 74). However, in one sense, Watkins falls into the opposite extreme to that of Langton: in Watkins's opinion, the special laws of nature depend on the nature of things in themselves, and hence phenomenal causal powers supervene on things in themselves, *tout court*, as it were. If this were the case, there would be a window into the world of things in themselves, contrary to Kant's repeated claims. In my view, by contrast, this dependency concerns only a subset of things in themselves—namely rational beings—and this assumption is required when assuming the freedom of the rational agent. To be sure, although for Watkins the dependency is global, this does not mean that freedom is a property of all things in themselves. Rational beings have the capacity to choose their nature, as he puts it, unlike non-rational things in themselves. A more detailed discussion of these issues will also be provided on another occasion.

and evidently this is the core of Kant's strategy, so that a reading in terms of supervenience can at best count as a rational reconstruction.

To counter this objection, we need to realize that there is a second function of the essential properties, and for that matter, of *individual* essential properties. They are the ultimate hub of all properties (i. e., the so-called "substratum"), including causal powers. To be sure, this is true of all substances in the world of appearances, but in so far as the hub-function of and supervenience on essential properties come together, it makes sense to speak of a causality of reason. Nonetheless, this causality is not a second type in the sense of being a causality besides natural causality; rather, it is the full picture of a rational agent's natural causality, situated in the architecture of being. In this picture the individual essence of the rational agent, or as it has traditionally been called, its substantial form (or rational soul), is what counts as a free cause. Strikingly, in the passage quoted above (5:94 f), what is "free" for Kant is not an action, but "I", and, at least within the framework of practical reason—in particular of the postulates of pure practical reason—this "I" is a substance.

With this in mind, we can finally return to the argument of the *Kritik der praktischen Vernunft*. What Kant has established now is basically a second conception of powerlessness or being in one's power, respectively; at least he has established one component of this conception. We have the conception of a power by virtue of a capacity to determine natural causal capacities, or to be more precise, to determine the specific constitution of these natural causal capacities. This is already enough to see how rationality, as something non-natural, can make a difference within the natural world. It does that by being both the substratum (i. e., the substantial base) and the supervenience base of natural causality. In this vein, the rational agent is the substantial form that functions as this substrate. To repeat, according to this analysis there is no second type of causal capacity involved; rather, one and the same causal capacity is anchored in different depths of the cause. Taken as a natural cause, the causal capacity is the capacity of a phenomenal substance; taken as a free cause, it is the capacity of a noumenal substance or substantial form.

In order to undercut the argument against freedom and to establish the second component for the second kind of power, we need to establish that, in a sense, the agent as a substantial form has options as to which specific constitution it "gives" these natural causal capacities. If this can be done, and Kant obviously thinks it is possible, then it is easy to see how the argument against freedom collapses, as the formal-



ization of Kant's argument developed above reveals: On reading A, even if we concede in the (modified) first premise that we are powerless in the stronger sense ("PL<sub>2</sub>p(t<sub>1</sub>)") with regard to the past, TNP and premise 2, namely  $N(p(t_1) \rightarrow p(t_2))$ , cannot turn this into powerlessness in this stronger sense about the future. On reading B, the inference from natural necessity to powerlessness in the stronger sense is blocked, because  $Np(t_2) \rightarrow PL_2p(t_2)$  in the (modified) fourth step is false. Whichever the correct reading may be, the underlying idea is the same: since natural necessity is relative to the actually existing set of special causal laws, if we have the capacity to influence what these laws are, then, of course, the transfer principles cannot transform our powerlessness insofar as we are appearances into powerlessness *tout court*. That said, Kant must also show that it is safe to assume that the agent as a substantial form is not in the grip of other potentially necessitating factors. What comes to mind as an example of such a potentially necessitating factor is God's creative activity, and this is precisely what Kant is discussing in the third part of the section under consideration. For him, TR fails to provide the metaphysical framework, within which the constitution of an individual agent as a substantial form can count as a bedrock fact. Which assumptions enter the TI position to make this possible will however have to be the topic of a different investigation.

PART III

The Role of Autonomy  
in Unifying Personhood



## 12. The Cognitive Dimension of Freedom as Autonomy

*Rainer Enskat*

### 1. Exposition of the Problem

As is well known, Kant characterizes freedom under many different aspects. One of the important questions guiding the judgment about the internal make-up of his theory of freedom is whether all these aspects and characterizations belong to a coherent and well-founded conception or not. In what follows I shall first examine one of the most prominent of these aspects and characterizations.

The guiding aspects and the corresponding characterizations belong, of course, to the practical dimension. Nevertheless, Kant develops a preliminary formal analysis of the concept of freedom in the Transcendental Dialectic of the *Critique of Pure Reason*, splitting this concept into a negative and a positive component. According to the negative component, to be free implies that a free being is *not* dependent on any sensible conditions (3:562, 585, 831).<sup>1</sup> This negative component corresponds to a certain part of our every-day grammar for speaking about freedom:—when we say a person is free *of* something (e. g., *of* a burden, *of* guilt, *of* deceases, etc.). Correspondingly, the positive component is in harmony with a certain other part of our every-day grammar for speaking about freedom:—when we say a person is free *for* something or free *to do* something or free to do something *in a certain way* (cf. 3:562, 713).

As Kant develops this formal analysis of the practical concept of freedom in a strictly theoretical context, he argues with emphasis that, with the help of theoretical means alone, deciding whether freedom is a fact or an illusion is not possible at all. Within the limits of the theoretical dimension the concept of freedom is condemned to stay, as he puts it, purely problematic (3:830–1). Consequently, this problematic status

---

<sup>1</sup> English quotations of Kant's words are my translations, from the *Akademie-Ausgabe*.

is the systematic starting point for his quest to find out how we can recognize or discover that freedom is a fact and neither a mere problem nor a mere illusion.

The result of Kant's search is well known. In the end he is convinced he has found the only reliable cognitive medium for the discovery of the fact of freedom. He circumscribes this medium in the first footnote of the Preface to the *Critique of the Practical Reason*, as the reason for recognizing the fact of freedom: in the Latin language of the scholarly tradition of his time he speaks of the *ratio cognoscendi* of freedom (5:4n). This medium for recognizing the fact of freedom he identifies with the categorical imperative. In this essay I will argue partially that this fundamental, freedom-manifesting function of the categorical imperative depends on another, a more specific cognitive function of the Categorical Imperative, one that Kant occasionally, in his lectures on ethics, characterizes as having the function of a principle of judgment (cf. 27, 2.2:1428) (i. e., as being a criterion); of course, as is well known, this is the criterion of morality.

Nevertheless, this cognitive medium of freedom is even more complex. Not only is the categorical imperative (i. e., the principle of judgment of morality) the reason for recognizing the fact of freedom. As Kant argues in the same footnote (5:4n), freedom is, conversely, the *ratio essendi* of the categorical imperative (i. e., the reason for the categorical imperative *to be* the principle of our moral judgments). Therefore, I am going to argue that freedom, according to Kant, *is nothing else* than (1) the cognitive faculty to judge and to recognize the moral character of maxims and of ways of acting, and at the same time, (2) the practical faculty to act *according* to such judgments and recognitions.

## 2. The Conditional Function of Freedom

When we concentrate on the cognitive functions of the categorical imperative we should, of course, never neglect the crucial fact that these functions serve exclusively within a strict practical context. That implies that these cognitive functions can be exercised exclusively in favor of two practical insights: *that* freedom is a fact and neither a mere problem nor a mere illusion; and *what* counts as a moral or a morally-consistent character.

In the practical sense freedom is, according to Kant, primarily *freedom of the will*, because the will has *actions* or *ways of acting* as its primary

object or content. In the same practical sense, but along the lines of the second insight, he characterizes freedom as *a practical form of causality*. That means a being that has a free will can, by the very help of his or her free will—and only by this help—be the cause of a special type of *practical* effects, namely of the moral *character* of the maxims of the will (i. e., more precisely, of the object or content of the maxims of the will), and this means (even more precisely) of the moral *character* of the ways of acting, for these actually *are* the objects or contents of the maxims of the will.<sup>2</sup> But, as this content of the maxims of the free will is identical with ways of acting, the most important practical effect that can be caused by the use of the free will is the moral character of ways of acting, not of the maxims as such or of the will as such.

To analyze Kant's concept of a cause in this context is not superfluous. This can be done in a very simple way if we pay attention to the concrete paradigmatic cases of causality that Kant presents in the context of his theoretical philosophy. One of these examples is the causal case of the sun melting a portion of wax. Of course, Kant knew well that a hot stove also melts a portion of wax (cf. 3:793). But such comparisons of simple causal cases can show that, according to Kant, a cause has a definite *conditional* status and function: it is *a sufficient condition* for what is effected by it. In the following I shall therefore adopt the premise that a being with a free will that makes use of its free will is, by this very use, a cause in the sense of being the *sufficient condition* of the moral character of such ways of acting.

At this point, I can state the thesis I will defend: freedom *consists*, according to Kant, *primarily* in the *cognitive faculty* to judge and to recognize and, *secondarily*, in the *practical faculty* of so-called causality of the free will, to exercise moral and morally-consistent ways of acting. To this complex—cognitive as well as practical—faculty Kant has given the name autonomy. But how does the crucial cognitive procedure of this autonomous practical judgment and recognition really work, as guided by the categorical imperative?

---

2 It was an extremely important insight of W. D. Ross, *Kant's Ethical Theory* (Oxford: Clarendon Press, 1954), 31–3, that the categorical imperative is a criterion to differentiate not different *actions*, but different *characters* of one and the same action (i. e., its moral and its non-moral or immoral characters). The neglect of this point is one of the most important sources of deeply misunderstanding Kant's ethical theory as a theory of freedom of choice (*Freiheit der Willkür, Wahlfreiheit*) between alternative *actions* – beginning with Fichte, Schelling, and Hegel.

### 3. The Procedural Depth-structure of the Categorical Imperative

If we want to reconstruct the single acts of this procedure in detail we should keep in mind one item that is decisive for the freedom-manifesting role of the categorical imperative. This item lies in the simple fact that the categorical imperative is intrinsically attached to a procedure of judging and recognizing the moral or the morally-consistent character of maxims and of ways of acting, a procedure that can and must be controlled exclusively under formal aspects. The formal character of this procedure is the most important manifestation of what Kant occasionally calls the “formalism of reason” (cf. 18:40, R4953). For this reason, of course, he characterizes the categorical imperative not only as a principle of judgment, but also as “The logical principle [of morality]” (cf. 29.1:621). But the (logical) formalism of this procedure is decisive for the freedom-manifesting function of the categorical imperative that guides this procedure in a double sense, according to the negative and positive components of the concept of freedom. According to the negative component, the formalism of the procedure guarantees that itself and its criteria are totally independent of any material, not to mention sensible elements; according to the positive component, this formalism guarantees that the judgment and the recognition of the moral or the morally-consistent character of a maxim or of a way of acting can be gained exclusively by formal criteria. This implies, from the very beginning, that freedom as autonomy includes a very special type of cognitive faculty: it is primarily the faculty to judge and to recognize the moral or the morally-consistent character of maxims or of ways of acting under exclusively formal aspects and criteria, and, therefore, totally independently of any material aspects and criteria. Therefore, in his theory of the categorical imperative Kant reconstructs the formal structure of the cognitive acts that enable an autonomous (i. e., a reason-guided) being to judge and recognize spontaneously (i. e., *without* having reconstructed this formal structure) his morally-consistent maxims or ways of acting.

#### 4. How to Apply the Categorical Imperative (I)

We start our analysis of the procedure best in the conventional way by taking verbally Kant's classical formulation of the categorical imperative in §7 of the *Critique of Practical Reason*: "Act in such a way that the maxim of your will can be valid always at the same time as principle of a universal legislation." (cf. 5:30). The first formal act of the judging procedure guided by this principle of judgment (i. e., by this criterion) concerns the formal structure of a maxim. A maxim is characterized by Kant as a subjective principle (cf. 5:19). The appropriate linguistic expression of the subjective form of such a principle is, of course, the first person pronoun (cf. e. g., 4:402, 403, 422, 423, 429, 438). The complete linguistic expression of the form of a maxim of the will, together with the intended way of acting, seems to be this:

(0) *I intend to act so-and-so.*

But, if we look closely to the concrete cases of such maxims, discussed by Kant, we can find that the formula (0) is not quite complete. Kant's discussions show that maxims have a structural appropriateness to certain types of practical situations where they are exercised. Therefore, we should represent this situative appropriateness in the complete linguistic representation of the form of the maxim. I propose the following representation:

(1) *I intend to act so-and-so in situations of type S\*.*

The next step in our analysis concerns a procedural aspect that has clearly dominated discussions for many decades—the generalization or universalization of a maxim. Generalization or universalization is a simple logical operation whereby an individual element of the logical material is transformed into a general or universal element. In the maxim the individual element is represented by the first person pronoun, so that the generalization or universalization transforms the maxim into a general or universal sentence of the will:

(2) *Everybody intends to act so-and-so in situations of type S\*.*

A very simple look should clearly indicate that this logical operation of generalization or universalization, or the result of this operation, is not of the slightest direct relevance for the moral character of a maxim or of a way of acting. The only relevance of this step lies in the fact that it makes manifest, by purely formal means, two further components of



that dimension of freedom corresponding to the negative and positive components of the concept of freedom. By the simple formal operation of generalization or universalization, the subject of the maxim shows that he or she is not, so to say, a prisoner of an egocentric or even solipsistic perspective, but is free *for* a universal perspective on all other possible subjects of practical maxims.

Nevertheless, even under purely procedural and formalistic aspects, this operation is by far not the last formal operation in the whole procedure of judging and recognizing. This can be seen by a simple look on Kant's formulation of the categorical imperative. For what is missing up to now is the act Kant points to by the condition that the maxim must be able to serve as a principle of general or universal *legislation*. I will give the name of *nomologization* to the corresponding formal act within the procedure.

Before explaining the linguistic expression of this nomologization I propose a formal simplification. We can see from the whole procedure that we can neglect up to the very end the volitional or intentional component of the maxim, so that we can concentrate on the practical content of the will (i. e., on the way of acting). With this simplification in mind the nomologization has to take account of the fact that a general or universal legislation in the practical field is, for Kant, a general or universal *obligation*. Therefore, the formula as revised for the result of the nomologization-act is this:

(3) Everybody *should be / must be / is obliged* to act so-and-so in situations of type S\*.

Nevertheless, closer inspection can show, at once, that this step of nomologization or universal obligation is only the first half of a double act. This double act must be completed by the act Kant has in mind when he speaks so emphatically of autonomy or *Selbstgesetzgebung* (cf. 4:431, 434 f) (i. e., of self-legislation and, by that, of self-obligation). Our closer inspection can show that this act, as far as its formal structure is concerned, is a *reflexive*, purely cognitive act whereby the original subject of the maxim *recognizes* and *acknowledges oneself* as one individual subject among all of those subjects whom he conceives as subjects of the universal obligation exercised by the preceding act (3) of the procedure of judging and recognizing. The linguistic expression of this reflexive act of recognition and acknowledgement, therefore, is this:

(4) I *should be / must be / am obliged* to act so-and-so in situations of type S\*.

I cannot stress strongly enough that this act is exclusively a purely cognitive act of recognition and acknowledgement; it is a subtle cognitive act, because it is a reflexive act of *self*-recognition and *self*-acknowledgement and at the same time an act of practical insight because the subject of the maxim recognizes and acknowledges himself as subject of a nomological self-obligation to act in the way characterized by the content of his or her maxim.<sup>3</sup> Of course, this very *reflexive* character of the *nomological self*-obligation makes practical autonomy a true *auto*-nomy and a true *auto*-nomy.

But I cannot less strongly stress that this reflexive act has the form of a *multiple identification*, though not simply of the twofold identification in the sense that (1) the subject of the maxim identifies oneself with (2) one of the *passive* subjects of a nomological obligation. Within the whole judging-procedure of this principle of judgment the subject of the maxim identifies himself or herself with (3) one of the addresses of the categorical imperative and with (4) the subject of the universalization or generalization and with (5) the subject of the nomologization. Even more than that, by the very act of executing this formal procedure of self-judgment, the subject of the maxim identifies himself or herself also with (6) the very subject of this whole procedure. Therefore, it becomes clear, by a closer look, that this multiple, six-fold reflexive identification is the formal substructure of the *practical* “unity of human personhood”. This unity is evidently a unity of a special multiplicity—the unity of *being* one and the same subject of many different cognitive, but practically relevant acts, *as well* as the unity of being *conscious* to be one and the same subject of this multiplicity of these acts. But this unity of

---

3 This *reflexive* act whereby the subject of a maxim recognizes and acknowledges *himself/herself* as one of all nomologically obliged subjects of the same maxim is evidently *presupposed* by the *non*-reflexive *social* act of acknowledgement that Hegel treats as basic in the *Phenomenology of Mind*, section B, a treatment that is partly adopted by contemporary Practical Philosophy. Cf. Ludwig Siep, *Anerkennung als Prinzip der praktischen Philosophie. Untersuchungen zu Hegels Jenaer Philosophie des Geistes* (Freiburg/München: Alber Verlag, 1979), esp. 131–45, and Axel Honneth, *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte* (Frankfurt/Main: Suhrkamp Verlag, 1992), esp. 274. The reflexive act of recognition and acknowledgement is conditional for not confusing one’s *ontological status* as a practically autonomous being with being an *adopter* of one (or more than one) *socially* created and acknowledge *role*; the social act of acknowledgement is conditional only for *acting with social success*, but not for being an autonomous subject (i. e., an autonomously *judging* subject).

being conscious of this identity is *achieved* by these reflexive acts of identification.

## 5. The Anthropological Dimension of the Categorical Imperative

With this reflexive cognitive act of self-obligation—and this means: with this autonomous act—by the original subject of the maxim, we have come to the end of the first half of the procedure of judging and recognizing. Quite clearly, this part is totally devoid of any moral-specific components, aspects, and criteria. We have to ask, therefore, how such specific elements can be brought into this procedure by Kant. To answer this question I want to draw attention to two passages in Kant's writings that are generally neglected, although they contain important material for such an answer, because they inform us about the human-specific, the specific anthropological aspect, of Kant's conception of morality.

The first passage is from his *Anthropology from a Pragmatic Point of View*. Kant's reflection in this passage is a type of so-called thought-experiment and has the logical form of a counterfactual. Kant presupposes

that on some other planet there might exist rational beings, who could not think in other ways than aloud, i. e., in waking as well as in dreaming, be they in company or alone, they might not be able to have thoughts which they did not utter at the same time. What curious kind of reciprocal behavior, different from our human species, would this generate? (cf. 7:332).

The most important answer to Kant's own question, especially in an ethical context, would be: such beings would not be able, by their very genetic nature, to lie. But Kant here does not argue in an ethical, but in morally neutral anthropological context. Therefore, we have to formulate the answer to Kant's question in a moral-indifferent way, such as the following. Such beings would not be able, by their very genetic nature, to hold silently for true the contrary of what they assert aloud to be true in communication.

But what is decisive in this anthropological context is not Kant's para-anthropological thought-experiment, but its central anthropological presupposition: all members of the human species are endowed, by their very genetic nature, with the dispositional faculty to deceive one another by uttering the contrary of what they believe to be true. This

faculty of deceiving is, as such, a purely technical faculty for deceitful communicative behavior. Kant himself discusses the whole point explicitly and exclusively under behaviorist aspects. This theory of a purely technical faculty and a purely communicative technique, is totally indifferent, morally. Therefore, what is decisive in this context is the fact that Kant's ethical theory of the central moral relevance of lying has a morally-indifferent presupposition in his anthropology of the human-specific faculties. I want to show that it is possible to build a bridge from this specific anthropological correlate to a specifically Kantian ethics, so that the latter no longer gives the impression of having an extremely narrow field of validity and application

## 6. Two Additional Anthropological Premises

The most important missing elements of the bridge we are looking for can be taken from the anthropological reflections of Günther Patzig and Edward Craig. Patzig has worked for nearly four decades on problems of ethics and of the so-called applied ethics, especially on Kant's ethics and on utilitarian ethics. By the anthropological reflection I have in mind, Patzig calls attention to a trivial truth that, nevertheless, will serve an important argumentative function in my reconstruction of the anthropological presupposition of Kant's ethics. Patzig mentions the fact that human beings are "beings who are in need of substantially appropriate information according to the reality in which they exist".<sup>4</sup> The second reflection comes from Craig's German-written book on *What we can know*.<sup>5</sup> He reminds us, strictly following the line of Patzig's anthropological argument, of the importance of not neglecting the fact that not a single circumstance of the reality we exist in can appear to be too insignificant to some person, in order that some other person is in need of being informed by somebody about exactly this very circumstance. Craig points, for example, to those circumstances that for each of us in any situation are hidden behind our back.<sup>6</sup>

---

4 Cf. Günther Patzig, *Wertrelativismus und ärztliche Ethik* (1988), republished in Günther Patzig, *Gesammelte Schriften II. Angewandte Ethik* (Göttingen: Wallstein Verlag, 1993), 54–72, here translating page 57.

5 Edward Craig, *Was wir wissen können. Pragmatische Untersuchungen zum Wissensbegriff* (Frankfurt/Main: Suhrkamp Verlag, 1993).

6 Cf. Craig, 88–9.

With these two anthropological premises in mind we can turn back to Kant's theory. I want to draw attention to two theses of Kant's that combine directly the anthropological aspect with the moral aspect. Both theses are from his late essay, *Announcement of the Near Conclusion of a Treaty for Eternal Peace in Philosophy* (8:411–22). The first thesis is: "The lie ... is the substantial fault in human nature"; the second thesis speaks of "the father of lying by whom all evil has come into the world" (8:422). If we combine these two theses with the anthropological theses of Patzig and Craig, we have the complete set of premises that are sufficient and necessary to make plausible the central anthropological pre-suppositions of Kantian ethics. The argument to make this plausible is this: if each human being, by his or her basic and life-long cognitive situation, is in need of appropriate information about the reality he or she lives in and if no human beings can be absolutely sure in any daily situations that they are fully informed about all relevant circumstances in any given situation, then each person must rely on the continuing informational veracity of any communicative partner.

This is the tacit plausibility-argument that must guide an ethic of the Kantian type when it concentrates on the moral case of lying. But, of course, it is not the crucial argument of this elaborated theory. This genuine Kantian argument must—and can—be reconstructed exclusively along the lines foreshadowed by the categorical imperative in order to judge, to recognize and to practice moral or morally-consistent characters of ways of acting. Of this procedure I have, up to now, presented only the first half.

## 7. How to Apply the Categorical Imperative (II)

When we start to reconstruct the second half we have to keep in mind that the first half should, by its formal character, be relevant for morally judging *any* maxim or *any* way of acting and, therefore, also for judging the maxim of lying or the act of lying. Therefore, we must take care of this concrete relevance, and we can do this along the lines of this reconstruction by applying the first four formulas directly to the case of lying. This can be done in our context in a very simple, technical way by substituting words for the act of lying:

- (1.1) I intend to *lie* in situations of type S\*
- (2.1) Everybody intends to *lie* in situations of type S\*

(3.1) Everybody should be / must be / is obliged to *lie* in situations of type S\*

(4.1) I should be / must be / am obliged to *lie* in situations of type S\*

The rest of the reconstruction may be guided exclusively by substantial reflections of Kant himself as he develops them in the context of a concrete case-study of lying in his *Foundations of Metaphysics of Morals*. With regard to the formal acts of judging that lead from (1.1) to (4.1) he asks (4:403): “Should I be content, that my maxim ... should be valid as a universal law (as well for me as for others)?” Examining this question, Kant continues: “I shall soon recognize that I can will the lie, but a universal law of lying I cannot will at all”. The argument Kant uses to show that this is a real practical insight has two parts. The first part argues that “according to such [a law] there could not properly be [a lie]”; the second part gives the reason why under such a law there could not properly be a lie: “... it would be *in vain* ... to pretend others, who would not, yet, believe this pretention”.

The substance of this argument is this:

- (1) Each liar presupposes—in favor of the success of his or her practical intentions, disguised by his or her lying—that there is at least somebody who relies on the veracity of his or her deceiving communication.
- (2) If all people are legally obliged to lie and all people are respectful to the corresponding law, there is nobody who can rely on the veracity of any communicative partner.
- (3) Both premises show that—under the universal obligation to lie—the normal success-presupposition of the liar is not fulfilled.

The formal nucleus of this structure can easily be represented by two sentences. The first shows the success-presupposition of the liar or of the lying-maxim:

(1.1.1) There is *somebody* who relies on the veracity of *me* as a liar;

The second shows the consequence from the universal obligation to lie:

(3.1.1) There is *not* somebody who relies on the veracity of any liar

These sentences obviously contradict one another. This form of contradiction is a well-formed type of that notorious contradiction that, as Kant asserts (cf. 4:431, 434), causes the will of the liar to entangle itself. But, as we now can see, the will entangles itself into this contradiction because of the way of acting that is its object or content, or the content of its maxim, not because of any mystical property of the will as such.

This contradiction is a purely formal property of the purely logical relation between the normal success-presupposition of the subject of the lying-maxim and the consequence of the universal obligation to lie. In a similar sense it is a purely formal property to be the success-presupposition of whatsoever, and it is as well a purely formal property to be the consequence of whatsoever. This shows that the subject of the maxim who is as well the subject of the whole procedure of judging his or her maxim is, in the course of this complex judgment, totally free in the negative sense (i. e., independent of all material aspects and criteria). But the subject is also free in the positive sense, because he or she is free to come, by this very formal procedure of judgment, to a practical insight—to the insight, as Patzig once put it, that it is discreditable for a rational being to intend or to practice a way of acting that suffers of a contradiction<sup>7</sup> (i. e., as I have argued, a contradiction regarding its success-presuppositions).

Now, it is well known that a contradiction is merely a superficial indicator, though the most important formal indicator, for a mistake in the depth-structure of the field where it appears. We have, therefore, to ask: of which concrete type is the mistake in the depth-structure of the practical field where the foregoing contradiction appears as an indicator?

## 8. Conclusions

The mistake in the depth-structure is presented by the semantic or propositional content of the contradiction. This mistake has the structure of a violation of the universal mutual veracity and thereby of the universal mutual reliability that each human being, by his or her basic informational situation, is always in need of in any life situation. This mistake can be avoided only by universal mutual veracity and thereby mutual reliability. Veracity must be, therefore, according to Kantian ethics, the central moral character of human ways of acting—either directly, when we communicate utterances with informational content, or

---

7 Cf. Günther Patzig, *Die Begründbarkeit moralischer Forderungen* (1967), republished in: Günther Patzig, *Gesammelte Schriften I. Grundlagen der Ethik* (Göttingen: Wallstein Verlag, 1994), 44–71, here translating page 67.

indirectly, in the sense that our ways of acting have to be consistent with this moral character.<sup>8</sup>

This indirect sense is decisive when we want to find out whether the range of validity and application of Kantian ethics is confined to the simple case of lying or not. By help of our analysis we can easily show that it has, indeed, a wide range. We can put the substantial argument very briefly: besides the manifest case of lying, each way of acting that has a harmful intentional disinformation for any interactive partner as its success-presupposition is inconsistent with the moral character of veracity. As can easily be seen, all culpable cases of the western and the western-inspired penal codes and of the corresponding codes of contractual rights are cases that have such a success-presupposition of non-veracity. This is why all such ways of acting are immoral within the range of validity and application of Kantian ethics. And this is the result of the purely formal procedure for moral judgment and recognition, foreshadowed by the categorical imperative and partly exercised by Kant himself in his case-studies.

---

8 The formal and the practical structure of the categorical imperative are occasionally compared to that of the Golden Rule: “Do to others what you would like to be done to you” (positive version) or “Do not do to others what you do not want done to you” (negative version, also attributed to Kong-zi/Confucius). It seems at first glance as if the Golden Rule exposes a *formal structure*, namely reciprocity between the acting subjects, that gives to it a *practical superiority* over the categorical imperative, since the latter does not expose, at least superficially, such a form of practical reciprocity. But superficial impressions can deceive heavily, especially in regard to the practical field of ethics: the mutuality of veracity and reliability is that formal structure of the most important *outcome* of the *application* of the categorical imperative whereby this outcome shows, at least, a certain prima-facie resemblance to the reciprocal practical structure highlighted *directly* by the different versions of the Golden Rule. Nevertheless, this formal prima-facie resemblance should not mislead us into the error of viewing these two forms of reciprocity as potentially isomorphic, though on different levels of reflection—not to mention the error of viewing the Golden Rule, in either version, as potentially having, by its superficial indication of practical reciprocity, any practical or conceptual superiority over the Categorical Imperative, with its somewhat hidden reciprocity of veracity and reliability. The Golden Rule makes an *individual* practical wish, intention or volition (“what *you* want or what *you* do not want”) *conditional* for the praxis of *all other* actors; the Categorical Imperative makes the *nomological* compatibility of *any* individual maxim with the *anthropologically*-based dependence of *each* human actor on reality-conforming and situation-adequate information by *any* other actor conditional for practicing this individual maxim.



From all this, we can easily see which type of freedom Kant envisages when he attributes to the categorical imperative the double cognitive function of a principle of judgment, and of a reason for recognizing (the fact of) freedom. Freedom, then, is neither a simple type of freedom of the will nor a simple type of freedom of acting. Rather, it is the freedom to judge and to recognize and finally to imprint, by the so-called causality of the will, the singular character of veracity on our maxims and on our ways of acting—directly on our communicative speech-acts and indirectly on the success-presuppositions of all other ways of acting.

## 13. Respect for Persons as the Unifying Moral Ideal

*Makoto Suzuki*

### 1. Introduction

As David Ross points out, there seem to be various duties, such as duties of fidelity, reparation, gratitude, justice, beneficence, self-improvement, and nonmaleficence.<sup>1</sup> What makes all of commonsensical duties moral, and if they are true duties, what explains why they are true? Ross provides no answer to this question and has been criticized for that.

Robert Audi and Mark Timmons suggest that the intrinsic end formulation of Kant's categorical imperative—respect for persons as ends—is such a unifying ideal that renders commonsensical duties “intelligible and even expectable”.<sup>2</sup> This is an interesting proposal not only for Rossians and Kantians but also for other moral philosophers.

I argue, however, that respect for persons can neither ground nor even unify commonsensical moral duties. The partiality of commonsensical obligations does not sit well with respect for persons. The proposal also faces a dilemma: if respect for persons is given a specific Kantian meaning, some commonsensical duties are not shown to be moral and true; if respect for persons is given a more broad and intuitive understanding, the ideal becomes too thin to explain why any commonsensical duty is moral and true. Thus, the intrinsic end formulation of Kant's categorical imperative fails to ground or unify commonsensical moral obligations.

---

1 W. D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930).

2 Robert Audi, “Intuitionism, Pluralism, and the Foundations of Ethics”, in his *Moral Knowledge and Ethical Character* (New York: Oxford University Press, 1997), 48; Robert Audi, “A Kantian Intuitionism”, *Mind* 110 (2001), 601–35 (Reprinted as chapter 3 of his *The Good in the Right: A Theory of Intuition and Intrinsic Value* (Princeton University Press, 2004), esp. 618; Mark Timmons, *Moral Theory: An Introduction* (Lanham, MD: Rowman & Littlefield Publishers, Inc., 2002), 203–4.

## 2. Unconnected Heap Objection against Ross's Theory

Ross holds that there is a plurality of (*prima facie*) moral duties, such as duties of fidelity, reparation, gratitude, justice, beneficence, self-improvement, and nonmaleficence.<sup>3</sup> The list of these obligations sounds really commonsensical, that is, intuitive to nearly all competent moral agents. Ross thinks the contents of moral obligations are self-evident.

One traditional criticism against Ross' position is that these proposed duties "seem unconnected—a heap of duties with nothing that ties them together so that they can be clearly understood as representing moral requirements."<sup>4</sup> This unconnectedness leaves us wondering what renders all of these commonsensical requirements moral and, if they are true duties, explains why they are true. Timmons calls this problem "the unconnected heap problem".<sup>5</sup>

## 3. Audi's Proposal and Timmons' Gloss

Audi and Timmons argue that the second formulation of Kant's categorical imperative can make sense of and unify commonsensical moral obligations; this thereby solves the unconnected heap objection to Ross' commonsensical moral system. The second formulation of categorical imperative goes as follows (4:429): "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means." Kant takes this idea, respect for persons in themselves as ends (or more simply, respect for persons), to be an expression of the fundamental unconditional requirement of morality. According to Kant, humanity or rational nature has dignity (i. e., unconditional worth). Rational nature is constituted by autonomy or freedom of choice (4:429)—that is, a capacity to act freely on the basis of reason and independently of desires. This is why we should respect persons (i. e., autonomous beings) as ends.

Kant apparently holds that one specific system of duties can be derived from the requirement of respect for persons. However, Timmons argues and Audi suspects that the idea of treating persons as ends in themselves is too vague to derive one specific system. For example, re-

---

3 Ross, *The Right and the Good*, 21–2.

4 Timmons, *Moral Theory*, 203.

5 Timmons, *Moral Theory*, 203.

spect for persons does not seem to determine whether or not the pacifist is right in saying all forms of intentional killing are wrong.<sup>6</sup> Apparently, the idea of respect for persons is so vague that it favors neither the pacifist system of obligations nor the anti-pacifist system.<sup>7</sup> However, Audi argues that respect for persons can still serve to make sense of the commonsensical moral obligations Ross formulates.

Is it not plausible to hold that in lying, breaking promises, subjugating, torturing, and the like, one is using people merely as a means? And in keeping faith with people, acting benevolently toward them, and extending them justice, is one not treating them as ends, roughly in the sense of beings with intrinsic value (or whose experiences have intrinsic value)? The point is not that Ross's principles can be deduced from the categorical imperative ... rather, the intrinsic end formulation of the imperative expresses an ideal that renders the principle of duty intelligible or even expectable.<sup>8</sup>

Timmons agrees that this is a plausible suggestion: the thought that commonsensical obligations are moral requirements is made sense of by the idea of respect for persons; this is not because these obligations are derived from the idea of respect for persons, but because "we can view" them as an interpretation of that idea.<sup>9</sup> I take Timmons to mean that commonsensical obligations are an intelligible and even expectable specification of the idea of respect for persons.

Are commonsensical obligations really an intelligible and expectable specification of the idea of respect for persons? We will consider this issue below.

---

6 Timmons, *Moral Theory*, 181–2.

7 In elaborating the indeterminacy of "respect for persons", Timmons quotes the following passage of James Griffin: "Every moral theory has the notion of equal respect at its heart: regarding each person as, in some sense, on an equal footing with every other one. Different moral theories parlay this vague notion into different conceptions ... [M]oral theories are not simply derivations from these vague notions, because the notions are too vague to allow anything as tight as a derivation." James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Oxford University Press, 1986), 208.

8 Robert Audi, "Intuitionism, Pluralism, and the Foundations of Ethics", 48. See also Audi, "A Kantian Intuitionism", esp. 618, and Audi, *The Good in the Right*, ch. 4, esp. 144–5.

9 Timmons, *Moral Theory*, 203–4; see also Mark Timmons, "Toward a Sentimentalist Deontology (Comment on Greene)", in W. Sinnott-Armstrong (ed.) *Moral Psychology Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development* (Cambridge, Ma.: The MIT Press, 2008), 96 and 100.

#### 4. The Partiality of Commonsensical Obligations

Problems arise from at least two directions. First, it is difficult to understand how the partiality of some commonsensical obligations can be explained in terms of respect for persons as ends. Those who have read Bernard Williams' criticism of Kantian ethics<sup>10</sup> will anticipate this problem. Commonsensical obligations require partiality when an agent has a special relationship with someone, herself, her family, friends, fellow persons, coworkers, compatriots, and so on. For example, consider the obligation of beneficence. To make the partiality of commonsensical obligations explicit, consider the following scenario, inspired by William Godwin's famous example of Fénelon and the chambermaid.<sup>11</sup>

Suppose that a public benefactor and your mother are in a house on fire. You can take the public benefactor to be any stranger who would benefit and even save many people's lives if he survived—a Mother Teresa, a Gandhi, a President Obama, or a scientist who has recently discovered but not yet published the cure for HIV. Your mother is a normal person, who would not make a comparable contribution to people's good even if she survived. Only you are near the house, and if you go into the house, you can save only one of them; if you call and wait for rescue, they will both burn to death. Whom should you save?

Most of people think you should save your mother instead of the public benefactor. Commonsensical obligations favor those close to the agent. However, because your mother, the public benefactor, and the people who would be saved by the benefactor are all equally persons, it is difficult to see how respect for persons makes sense of requiring you to save your mother. Because your mother and the public benefactor are both equally persons, there seems to be no reason to save your mother over the benefactor. Moreover, because the benefactor would save many people while your mother would not, respect for person apparently requires that you save the benefactor over your mother: in that way, you can respect more persons.

This example is intended to illuminate the partiality that commonsensical obligations—in this case, the obligation of beneficence—in-

---

10 Bernard Williams, "Persons, Character, and Morality", in his *Moral Luck* (New York: Cambridge University Press, 1981 [973]), 1–19.

11 William Godwin, *Enquiry Concerning Political Justice, and Its Influence on General Virtue and Happiness* (1793), book 2, ch. 2: "Of Justice".

volve. The duties regarded as self-evident by Ross involve partiality; Ross emphasizes “the highly personal character of duty.”

The essential defect of the “ideal utilitarian” theory is that it ignores, or at least does not do full justice to, the highly personal character of duty. If the only duty is to produce the maximum of good, the question who is to have the good—*whether it is myself*, or my benefactor, or a person to whom I have made a promise to confer that good on him, *or a mere fellow man to whom I stand in no such special relation*—should make no difference to my having a duty to produce this good. But we are all in fact sure that it makes a vast difference.<sup>12</sup>

However, as we saw, because, apparently, respect for persons enjoins the equal treatment of each person, it is difficult to see how this idea can make full sense of the partiality of commonsensical obligations.

## 5. The Wider Scope of Commonsensical Obligations

The second problem arises from the scope of commonsensical morality. Commonsensical morality concerns not only human beings, but certain other living things. Mammals, birds, reptiles, amphibians, fishes, and perhaps others can feel pain, so it seems that we have obligations of non-maleficence and beneficence toward these animals. Because most of these animals are not persons, it is difficult to see how respect for persons can make sense of these obligations toward animals. Kant himself denies that we have any duty toward animals; we should avoid cruelty against animals just because it damages our kindly and humane qualities toward humans (27:458–60). However, it seems that hurting or benefitting nonhuman animals morally matters in itself: we owe duties of nonmaleficence and so on to them. Ross himself argues that we have duties concerning animals not for the sake of persons but for the sake of animals themselves:

On the other hand, if we think we ought to behave in a certain way to animals, it is out of consideration primarily for *their* feelings that we think we ought to behave so; we do not think of them merely as a practicing-ground for virtue. It is because we think their pain a bad thing that we think we should not gratuitously cause it.<sup>13</sup>

---

12 Ross, *The Right and the Good*, 22; italics added.

13 Ross, *The Right and the Good*, 49.

As the contemporary discussion of moral status reveals, the problem of scope does not stop with obligations toward nonhuman animals.<sup>14</sup> Human beings with severe mental disabilities fail to be persons. Babies and young children are only potential persons and not actual persons. This is because there are certain necessary conditions for you to be a person. Respect for persons is taken to be respect for the rational autonomy of persons, and people with severe mental disabilities, babies, and young children fail to have such a capacity. It is thus difficult to see how respect for persons can make sense of commonsensical obligations to mentally handicapped humans, babies, and young children.<sup>15</sup>

One way to respond to the second problem is relaxing the requirement for being a person. Certain ways of relaxing the requirement are arbitrary. Suppose, for example, you take being biologically human to be necessary and sufficient for being a person. Then, respect for persons—now it turns out to be respect for any human being in the biological sense—can make sense of obligations toward mentally handicapped humans, babies, and young children. It still fails to make sense of obligations toward nonhuman animals, but it comes closer to vindicating commonsense morality. However, this way of relaxing the requirement for being a person appears to be arbitrary. Suppose there were some alien or android who is as intelligent, self-conscious, and autonomous as a normal adult human being is (Yoda in *Star Wars* or Data in *Star Trek: The Next Generation* would be an example, if either existed). The alien or android is not a human being in the biological sense, so respect for human beings would not prevent us from harming (or breaking) them. It seems, however, that because that alien or android is as intelligent and autonomous as a normal adult human being is, this discriminative treatment is arbitrary. Because Kant counts certain nonhu-

---

14 See, for example, Mary Ann Warren, *Moral Status: Obligations to Persons and Other Living Things* (New York: Oxford University Press, 1997), 101–3. Warren quotes the following passage of Tom Regan: “All that can be said about our dealings with such humans [on Kant’s theory] is that our duties involving them are indirect duties to rational beings. Thus, I do no moral wrong to a child if I torture her for hours on end. The moral grounds for objecting to what I do must be looked for elsewhere—namely, in the effects doing this will have on my character.” Tom Regan, *The Case for Animal Rights* (Berkeley and Los Angeles: University of California Press, 1983), 182.

15 In addition, possible future generations are not persons now, and who will be born to be persons depends on how we—current generations—behave. So if respect for persons concerns currently existent persons only, there will never be obligations to future generations, as commonsensical morality suggests.

mans—God and angels—as persons, he would object to this way of distinction, too.

In order to avoid this sort of arbitrariness, we must distinguish persons from nonpersons in a morally relevant way. For example, if you take all and only sentient beings—individuals that can feel pain—to be persons, this will not be arbitrary. You can then perhaps make sense of obligations to severely mentally handicapped humans, babies, and young children as well as to certain nonhuman animals. However, the problem is that this way of relaxing the requirement of respect for persons cannot make better sense of the system of commonsensical obligations than of, say, utilitarianism (i. e., the obligation to promote everyone's good as far as possible). Utilitarianism involves one version of the idea that moral agents must respect all and only sentient beings as ends: the happiness of all and only sentient beings count equally. Utilitarianism, however, is incongruous with a Rossian understanding of commonsense moral obligations. Utilitarianism allows the violation of Rossian obligations when doing so would maximize the net happiness, where every sentient being's happiness is counted equally. Ross contends, however, that commonsensical obligations cannot be overridden by the mere increase of net happiness.

Someone might here object that the idea of respect for persons is incompatible with utilitarianism. One could argue that utilitarianism can justify use of force, coercion or deception, so it can fail to respect persons as ends: it can use them as mere means.<sup>16</sup> However, first, utilitarianism justifies use of force, coercion or deception only if it maximizes the total happiness, where the happiness of every sentient being is counted equally. It is unclear whether use of force, coercion or deception in such a case uses persons as mere means. Second, it is implausible to hold that respect for persons bans all uses of force, coercion or deception. Suppose, for example, force or deception is needed for the defense of oneself or another person against serious offense. In such a situation, is it the case that use of force or deception involves using some person as mere means and so it is unjustifiable? Given these two points, it seems that *if all (and only) sentient beings are persons*, the idea of respect for per-

---

16 E.g., Christine M. Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals", Tanner Lecture delivered at University of Michigan (2004); online text at [www.tannerlectures.utah.edu/lectures/documents/volume25/korsgaard\\_2005.pdf](http://www.tannerlectures.utah.edu/lectures/documents/volume25/korsgaard_2005.pdf), 3.



sons makes sense of utilitarianism as well as of the system of commonsensical obligations.

Thus, relaxing the requirement for being persons in a morally sensible way apparently fails to make sense of the system of commonsensical obligations any more than another moral system, such as utilitarianism. Put more generally, the problem is that if we relax the requirement for being persons in a non-arbitrary, plausible way, it is difficult to avoid making the idea of respect for persons compatible with a non-commonsensical moral system. And if respect for persons is compatible with a non-commonsensical moral system like utilitarianism, commonsensical obligations are not an intelligible and expectable specification of the idea of respect for persons.

Let me note that the major contemporary attempts to make sense of obligations toward nonhuman animals within the Kantian framework do *not* appeal to the idea of respect for persons. Allen W. Wood and Christine M. Korsgaard respectively try to show that Kant should have endorsed obligations toward nonhuman animals. However, they do not argue that the obligations are made sense of by the Formula of Humanity in itself. Wood argues that respect for rational nature requires not only respecting persons (i. e., the beings that actually have it), but also behaving with respect toward the beings who, like children and nonhuman animals, have rational nature only potentially or virtually, or who have had it in the past, or have parts of it or necessary conditions of it.<sup>17</sup> Korsgaard holds that nonhuman animals are not rational beings and hence not persons. However, she argues that we have obligations toward them because a concern for the good of animals is implicit in our endorsement of our self-concern, and this creates obligations.<sup>18</sup> Perhaps we should not expect that the idea of respect for persons makes sense of commonsensical obligations toward non-persons, when even the major defenders of Kantianism think otherwise.

---

17 Allen W. Wood, "Kant on Duties Regarding Nonrational Nature", *Proceedings of the Aristotelian Society Supplement* 72 (1998).

18 Korsgaard, "Fellow Creatures", esp. section 6.

## 6. Conclusion

Given the partiality and wider scope of commonsensical obligations, they are not an intelligible and expectable specification of the idea of respect for persons. Then, despite what Audi and Timmons argue, the idea of respect for persons *alone* does not make sense of the thought that these commonsensical obligations are true moral requirements. The idea of respect for persons is *not by itself* a unifying moral ideal; some additional factor, or some distinct or more fundamental ideal is required for making sense of the thought that commonsensical requirements are moral and true duties. The Unconnected Heap objection against Ross' view of commonsense morality is still alive and kicking. And Kant's successors are left with the task of accommodating or explaining away the partiality and wide scope of commonsense morality.<sup>19</sup>

---

19 The penultimate version of this essay was presented at the Kant in Asia conference, held in Hong Kong from May 20 to 23, 2009. I thank all the people who commented on that version at the conference.

## 14. Kant and Virtuous Action: A Case of Humanity

*Vasil Gluchman*

A classical formulation of the philosophical or moral philosophical understanding of the meaning of the concept of humanity was provided by Immanuel Kant, who wrote (6:462) that

humanity itself is a dignity, for man can be used by no one (neither by others nor even by himself) merely as a means, but must always be used at the same time as an end. And precisely therein consists his dignity (personality), whereby he raises himself above all other beings in the world, which are not men and can, accordingly, be used—consequently, above all things.<sup>1</sup>

At another place (6:434–5) Kant specified the meaning of the concept of dignity or human dignity and wrote that

man as a person, i. e., as the subject of morally-practical reason, is exalted above all price. For as such a one (*homo noumenon*) he is not to be valued merely as a means to the ends of other people, or even to his own ends, but is to be prized as an end in himself. This is to say, he possesses a dignity (an absolute inner worth) whereby he exacts the respect of all other rational beings in the world, can measure himself against each member of his species, and can esteem himself on a footing of equality with them.

The fact that his understanding of humanity and man is entirely anti-nature and anti-biological is also evidenced by his statement that “it is one’s duty to raise himself out of the crudity of his nature, out of his animality more and more to humanity, by which alone he is capable of setting himself ends” (6:387). Then, to save a strange man is virtuous action because we had to overcome more moral hindrances than in a case of familiar people. However, is a strange man a more human or rational being than the familiar person?

I will present two examples of the understanding of humanity in contemporary Kantian ethics: Christine Korsgaard and Marcia Baron. Korsgaard states that, according to Kant, to respect the humanity of others means to share their goals. Such a status of humanity is, in her opin-

---

<sup>1</sup> Immanuel Kant, *Ethical Philosophy*, tr. James W. Ellington (Indianapolis: Hackett, 1983).

ion, the source of normative requirements and the source of all value.<sup>2</sup> Baron believes that in contemporary Kantian ethics humanity means above all respect and love for others and oneself as well. While in relation to others it means to help them achieve their permissible goals and to respect others' lives and characters, in the case of humanity expressed in relation to oneself it reflects the effort to achieve self-perfection and the development of one's talent.<sup>3</sup>

The above-presented notions of humanity are characterized by very strong anthropocentrism, formulated from the ethical (Kant and Kantian ethics) standpoints that more or less clearly postulate the superiority of human kind over nature, or over the animal realm. Contemporary genetics, neuroscience, biology, zoology, ethology, etc., however, present ever-new knowledge of genetic similarity between humans and the animal realm; the similarity exists between human brain activity and behavior and the brain activity and behavior of many representatives of the animal realm, especially primates or mammals, but also some other kinds of animals. So, is humanity a specifically human quality that separates mankind from nature, whereby mankind overcomes animality (i. e., our biological and natural determination), as Kant expressed it? And what actually is humanity?

According to common sense morality, we often understand humanity on the one hand, as respect for and acceptance of human being, and on the other hand, as support of the effort to develop its strengths and abilities. Let us think about these individual aspects of humanity and decide to what extent it is really possible to perceive them as adequately expressing the meaning of the concept of humanity. On the common sense morality level humanity is first of all respect for human being. This means that in the case of others we respect their ontological or metaphysical status as human beings (i. e., that they are above all the bearers of the morphological signs belonging to human being, having physiological similarity with people). This results in the duty to behave toward them as members of the same species—that is, as beings that are equal to us. In the case of support of efforts at the development of someone's powers and abilities, this usually means the creation of economic, social, mental, cultural, intellectual, and educational conditions for this

---

2 C. M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 299.

3 M. Baron, "Kantian Ethics", in M. Baron, P. Pettit, and M. Slote (eds.), *Three Methods of Ethics* (Oxford: Blackwell, 1997), 28–31.

human being's achievement of full development. This, then means common sense morality is a totally anti-Kantian form of understanding humanity, because it concerns especially the human nature of our fellows, not an effort to overcome its crudity.

The initial thesis of our thinking about humanity (i. e., humanity as the respect for human life) also appears to be anti-Kantian because it is more a natural or biological than solely a moral factor. Undoubtedly, morality has its biological basis related above all to the value of human life.<sup>4</sup> Bekoff believes the origin of virtue and morality is older than the human species.<sup>5</sup> If we are to summarize the ideas developed so far, following life sciences knowledge about similarities in the behavior of human beings and many animals, then we have to state that a great majority of the manifestations of human behavior that we usually call humane has a mainly biological or natural dimension that we have in common with other animal species, especially mammals and primates. These manifestations of behavior include the ones related to the protection and maintenance of our own life, and the lives of our children, relatives, friends, and acquaintances. In relation to these people (if we respect and support their lives, interests, and goals) we behave essentially in the same way members of various animal species, especially mammals and primates, behave toward those close to them. Such behavior, despite the fact that it is very needed and desired, can in no case be seen as uniquely human, or as something that can create a basis for humankind's claim to a special status in relation to other mammals and primates. The basis of our behavior is biological or natural. People, nevertheless, ascribe a moral value to it, and it is debatable whether such behavior is an example of the crudity of human nature. I think we can ac-

---

4 M. Bekoff, "Wild Justice and Fair Play: Cooperation, Forgiveness, and Morality in Animals" *Biology and Philosophy* 19.4 (2004), 489–520; "Animal Passions and Bestly Virtues: Cognitive Ethology as the Unifying Science for Understanding the Subjective, Emotional, Empathic, and Moral Lives of Animals", *Zygon* 41.1 (2006), 71–104; M. Ridley, *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* (New York: Viking, 1996); M. Ridley, *The Cooperative Gene* (New York: The Free Press, 2001); M. Ruse, *Taking Darwin Seriously: A Naturalistic Approach to Philosophy* (Amherst, N. Y.: Prometheus, 1998); B. N. Waller, "What Rationality Adds to Animal Morality", *Biology and Philosophy* 12 (1997), 341–56; R. Wright, *The Moral Animal: Why We Are, the Way We Are: The New Science of Evolutionary Psychology* (New York: Pantheon Books, 1994).

5 Bekoff, "Wild Justice and Fair Play", 515–6.

cept it as a fundamental moral value of humanity, as respect for human life, despite the fact that it is naturally or biologically based.

However, we can see human behavior that is not biologically or naturally based, but has an exclusively moral basis, such as our behavior related to the protection of variously disabled forms of human life. Another example is our behavior toward strangers, if it is aimed at the protection and maintenance of their life, property, physical or mental integrity, and their goals, and if it intends to help protect and maintain life.

Furthermore, we can think about other aspects of humanity and on this basis define the active and passive forms of the realization of humanity. The active form means the direct involvement of moral agents through their participation in activity developed for the benefit of strangers in need of help. The passive form of the realization of humanity means that our behavior expresses sympathy with strangers affected by disaster. Especially in the case of the latter form, an important role is played by moral feelings. Usually this passive form of humanity forms the basis for, or is the condition for the realization of, the active form of humanity (i. e., for the providing of assistance to those people who need it). Sympathy with the suffering often (though not always) leads to acting for the benefit of these people. Of course, active help is always more valuable than mere sympathy, but we should not minimize the value or potential of the humanity comprising sympathy. Our capacity to forgive comprises a similar potential for being humane. Forgiveness, just like sympathy, can be the initial point for our further acting, for the active realization of humanity in the form of assistance to others. Passive humanity can also be reflected in not acting (i. e., not causing harm to other persons despite the fact that the moral agent could do it while realizing his rightful intentions and goals). This passive form of humanity can be seen as a minimal level of humanity, related to the fact that the moral agent who cannot help another person should express sympathy, or at least not act in a way that could harm the other in the realization of rightful intentions and aims. The active form of the realization of humanity can be divided into positive and negative. The positive form means a direct assistance to a stranger who needs it in the realization of positive intentions and goals. The negative form means to prevent another person from the realization of harmful aims, such as intentions that could adversely affect some stranger.

My thinking, developed so far, has brought me to the conclusion that it is possible to respect humanity from the metaphysical or ontological perspective (i. e., to perceive someone as a human being on the basis

of morphological and functional signs belonging to human beings); this means to respect their nature, to respect human life. A fundamental moral value of humanity consists in such respect. From the ethical or moral perspective, humanity has to be realized and not only respected because it implies acting to the benefit of strangers in need. It is also latently present in the passive form of humanity (i. e., in the feeling of sympathy with the suffering or misfortune of strangers or in the case of forgiving someone). This is an additional moral value of humanity and it can be realized only through our behavior and conduct in relation to strangers. Virtuous action is then only an active form of helping strangers with their needs, because it means to overcome some moral hindrances; however, it has to be in accordance with moral duty.

We now have to distinguish between the generic behavior and conduct of humans that, despite having a biological-natural basis, also comprises a positive moral dimension related to the protection and maintenance of human life, and the behavior that also has a biological-natural basis, but comprises a negative moral dimension. In the animal realm the protection and maintenance of life, on the one hand, or its destruction, on the other hand, have no moral dimension or effect. Both manifestations of behavior, protective and destructive, are the natural manifestations of animal behavior and do not evoke any wider reaction among the members of the animal realm. Their effect is temporary and their impact limited to the local area. In the case of human society the reaction to such behavior is wider and, owing to the media, can cross the local borders where such kinds of behavior happen very quickly, especially if the behavior represses or destroys human lives.

That is why I suggest we speak about humanity in all the cases where human life is protected and maintained: it brings positive consequences for human life, with the specification that if it is the protection and maintenance of one's own life, or the lives of our relatives, friends, or acquaintances, then this is humanity based on a biological-natural foundation that, however, also has its moral dimension and effect, as a fundamental moral value of humanity. On the contrary, the manifestations of the protection and maintenance of life in relation to strangers represent an additional moral value of humanity (perhaps, in contexts relating to duty, virtuous actions); that is, they are the results of our cultural evolution, our moral development. In this way we accept all the positive manifestations of our behavior in relation to other people. Especially, we emphasize the value of helping, the protection and maintenance of the handicapped forms of human life and of the strangers

who need it, because such behavior transcends our biological-natural dimension, the basis we have in common with many other representatives of the animal realm.<sup>6</sup> The additional moral value of humanity (especially active helping to strangers) is an expression Kant's understanding of humanity and our effort to overcome hindrances. We can call it a virtuous action corresponding to Kant's idea of virtue, if it follows a moral duty.

In the first case we understand humanity as a generic, natural-biological quality typical of the behavior of members of the human species (i. e., as a fundamental moral value), while in the second case we understand it as an additional moral quality (perhaps as virtuous action, if it is in accordance with moral duty) that, despite having features of similarity with the first quality, differs in respect to the object of its realization. Despite the fact that in the first case we understand humanity as a biological-natural quality of mankind, this understanding of humanity cannot be identified with the biological makeup of humanity, because my understanding of humanity is related only to behavior leading to the protection and maintenance of human life. If this understanding of humanity is not to be influenced by speciesism, then we have to accept that in the animal realm, especially in mammals and primates, the protection and maintenance of one's own life, the life of offspring and other members of the herd, flock, or pack is a natural-biological quality typical of their species, and that we can call it *animality* and see it as equal in its forms or manifestations of behavior to humanity as the primary, natural-biological quality typical of the human species.

To summarize our points, we can state that humanity is understood as all the forms of behavior leading to the protection and maintenance (i. e., respect and development) of human life. On the basis of the differences in the objects of our behavior and conduct, we distinguish between humanity as primary natural-biological quality (the fundamental moral value of respect for human life) and as an additional moral quality

---

6 Bruce N. Waller in this respect speaks about the morality of care and the morality of duty. In his opinion, human rationalistic morality is an improved animal morality of care. Ethics of care is in its essence valid because affection, care, trust and generosity form a moral basis. The attitude based on rational principles is an important means of the widening, improving and supporting of moral behaviour when affection reaches its limits. The moral basis of the morality of duty resides in care and affection. The affection is rooted in biology, supported by direct and indirect reciprocity and exists prior to rationality. The rational morality of duty is an adaptive complement of the morality based on affection and care (Waller, "What Rationality Adds to Animal Morality", 353–4).



(in some contexts, a virtuous action) supporting and developing the human life of strangers. The moral value of the first kind of behavior is determined by our biological or social relations to those close to us. In the second case, the moral value of our behavior to strangers is a pure manifestation of our morality and I think it is fully consistent with Kant's seeing humanity as overcoming our nature and moral hindrances. In the first case the protection and the maintenance of life is a result of our basic value orientation, including our moral values that result from this orientation. In the second case our behavior and conduct for the benefit of strangers brings additional moral value. The basic form of humanity (its fundamental moral value) resides, then, in the protection and maintenance of one's own life and the lives of our relatives, friends, and acquaintances. It is the alpha and omega of our behavior and creates the basic natural-biological framework for our morality. It also creates the foundation for the basic rights and duties related to the protection and maintenance of human life. I do not think it is an expression of the crudity of human nature. On the contrary, it is the first, and very important, step to preserve real human nature in human beings.

On the other hand, the protection and maintenance of the life of strangers is a moral additional value (perhaps, virtuous action) whereby we create a new, higher quality in our behavior in relation to other people. In this case we can really speak about humanity as a moral quality, or value in Kant's sense. It is something that really is specifically human and deserves respect and admiration. By such behavior humans prove they can, at least to certain extent, transcend the natural-biological framework of their determination. Especially in that context it is a very close to Kant's ideas on the extension of the moral realm to strange people.

This can be achieved through the moral principles and particular moral norms that define some ways of pursuing humanity in individuals and in the social life of moral communities. I do not think that humanity as a moral quality is an unachievable and abstract moral ideal that is too far from the practice of moral agents. I mean that humanity as a moral additional value is the expression of actual requirements and interests of the individuals and of humankind in general. Human beings hope for their rational existence and survival through the application of humanity, its principles and respect for human dignity. Human existence also depends on the solution of environmental issues, an external condition

for the preservation of human life in general. However, what is important is that the moral agent should try to act humanely in his life.

Humanity is one of the most significant moral principles that forms the basis of human society, as a society of co-operating individuals. We can see that the future of humankind is possible only if we accept and apply the principle of humanity as one of these fundamental principles. This is not only because it is a basic duty of mankind, toward the preservation of our future existence, but also because of the danger of environmental disaster. It is so urgent that it is necessary to overcome narrow anthropological views on the future of the world and its life.

The idea of the preservation of human existence must be associated with respect for and the application of humanity as a moral quality, for this is the only possible response to the future of humankind. International co-operation of states and nations is the means of fulfilling humanity in the ordinary life of individuals and in the whole human society. Such co-operation assumes a perspective of the preservation of human life. One of the most significant conclusions of this reasoning is the idea that the meaningful existence of moral agents, communities and the whole of humankind is possible only through acceptance and application of humanity. I do not think it is an abstract and unachievable goal for most people during their lives. To respect and apply humanity in our lives, we do not need to be the saints, or enemies to our friends, as Friedrich Schiller ironically wrote on Kant's moral psychology.<sup>7</sup> Being human is enough. That is why I think we can justify the attribute *human and moral being* by our actions, regardless the unfavorable character of the contemporary period that perhaps tends to stimulate the opposite position. Despite this, I think no other alternative than the acceptance and application of humanity in the world is possible. It will be very useful for the whole of humankind if (in contrast with Kant's ideas) helping others (not only relatives) becomes natural for us, apart from moral obligation.

By and large, we can accept in some measure Kant's preference of strangers because it can help to promote stability, security, and order. I conclude—as if in close connection with Kant's views—that, despite its numerous mistakes and flaws, mankind has indeed been increasing the number of objects of its moral concern and protection; that is, hu-

---

7 Friedrich Schiller, *Xenien*, tr. H. J. Paton in *The Categorical Imperative* (London: Hutchison, 1967), 48.

manity has been showing at least some moral progress.<sup>8</sup> I would therefore like to end with Tomáš Garrigue Masaryk's idea that could be an appeal to the future of the whole of humankind at the beginning of the twenty-first century: "[H]umanistic ideal, [authentic] humaneness, is the foundation of all strivings of our time—particularly those that prevail at present in our national life. It is this which Kollár means when he says: 'When you cry, Slav, may it always mean Man'".<sup>9</sup>

---

8 V. Gluchman, "Morality of the Past from the Present Point of View", in V. Gluchman (ed.), *Morality of the Past from the Present Point of View* (Newcastle: Cambridge Scholar Publishing, 2007), 1–10.

9 T. G. Masaryk, *Humanistic Ideals* (Lewisburg: Bucknell University Press, 1971), 61.

## 15. Freedom and Value in Kant's Practical Philosophy

*Adriano Naves de Brito*

### 1. Autonomy and Experience: The Compatibilistic Knot

The Gordian Knot of compatibilism is the harmony between natural causality and freedom of action. From a strictly empirical perspective, an action's motive—corresponding, on the practical level, to what causes are on the theoretical level—is an event, and as such must have another event as a cause behind it. Experience, in the terms Kant presented it to us in *CPR*, implies the subsumption of the phenomena to the rules of apprehension of sensitivity, space and time, and to the units of synthesis in judgments, including causality. From the theoretical standpoint action, like any other phenomenon, is inserted in the causal chain of events.

The Kantian distinction between phenomenon and thing-in-itself opens up an important space of indetermination for morality. Indeed, that an action is empirically inserted in the flow of causes, so that the event that will cause a given motive is *ab ovo* predetermined *in re* (considering the thing only as a phenomenon)—that it be so, I say, does not mean that the intelligibility of this action, if sought by reason in its practical (and therefore noumenal) use, corresponds to the sum of the events of nature. It only means that it is essential to the scheme of apprehension of what is real, that there be a determination for this motive. The intelligibility of action, in turn, belongs to another sphere in the Kantian architecture.

According to *CPR* (A538/B566):

Ich nenne dasjenige an einem Gegenstande der Sinne, was selbst nicht Erscheinung ist, *intelligibel*. Wenn demnach dasjenige, was in der Sinnenwelt als Erscheinung angesehen werden muß, an sich selbst auch ein Vermögen hat, welches kein Gegenstand der sinnlichen Anschauung ist, wodurch es aber doch die Ursache von Erscheinungen sein kann: so kann man die Kausalität dieses Wesens auf zwei Seiten betrachten, als *intelligible* nach ihrer Handlung, als eines Dinges an sich selbst, und als *sensibel*, nach den Wirkung derselben, als einer Erscheinung in der Sinnenwelt. Wir würden uns demnach von dem

Vermögen eines solchen Subjekts einen empirischen, imgleich auch einen intellektuellen Begriff seiner Kausalität machen, welche bei einer und derselben Wirkung zusammen sattfinden.

The transcendental turn typical of Kant's theoretical philosophy creates the conditions to resolve the issue of determinism, as fatal for the survival of free will, by creating a realm of intelligibility. However, it does not open up the field for freedom in an objective sense. After all, every action will also be the result of a motive—a cause in the temporal chain—and this occurs within the sphere of nature. Consequently, an action will be as free as the occurrence of some event A is casual. Once the cause of this and the motive of that have been discovered, casualness and freedom, respectively, will disappear (cf. B577). Limiting the use of the categories to the sphere of apprehension of the world by experience does not, therefore, solve the problem of assigning responsibility, if we make it depend on freedom. It only creates the problem of uncertainty. Everything is determined causally, but the specific causes must be experimentally discovered, and there are no certainties about the relationship between them and their effects. In this context, and thinking about human actions, one can give the concept “freedom” some content, but it will be negative and much to the taste of Humean empiricism: ignorance and uncertainty about the causes of a practical phenomenon.

In the strict sense of the compatibilists (but also in that of a radical incompatibilist, of the kind who advocates complete separation between the natural and moral worlds), to assign responsibility requires that the actions of the agent *may* be motivated by determinations that are exclusively intrinsic to the person's will. From this perspective, responsibility is a corollary of autonomy, and an action will only be free if the will of the agent, on performing it, inaugurates causal chains. If, in a given case, the will does not determine itself, then, although it could be free, the action that results from it under these conditions will not be so. This action will only be a further link in the causal chain of events.

## 2. Responsibility and Freedom

The assignment of responsibility always requires the identification of a specific action that it weighs on. Speaking of the general responsibility of will does not explain much. When judging moral character, specify-

ing the particular act that would make a given will be responsible is essential. In assigning responsibility, the easiest case is when the specific action is considered free. Then, by definition, the agent is responsible for it. In another particular case, when the action does not result from an act of free will—when, thus, it is not the work of will, but occurred because will did not intervene—is the assignment of responsibility compromised?

The problem is how to make the agent responsible for actions that are not *stricto sensu* one's own work, for actions one performed without deliberating freely about what to do. The answer to this problem is that the agent *should* have deliberated, since he or she *could* have done so. Yes, one could, but did not; and the result of the argument is that the agent is responsible also for actions that he or she does not perform freely, for actions that occur according to the causal flux of nature, for actions, finally, of a will that, in a given circumstance, does not act autonomously.

Would the position of our compatibilist interlocutor, at least as regards the actions performed while enjoying autonomy of will, be safe, or not? In fact, no. How can one know whether or not some moral agent acts according to an autonomous will? One might answer: by evaluating the action. But this can only be investigated if it can be the object of a possible experience. In the sphere of possible experience, however, there is no way of determining when and if an action is free, since the autonomous causality involved in the act cannot be apprehended empirically. For finite beings, like humans, it is not possible to distinguish the autonomous character of an action—at least, not by observing its occurrence in nature.

Where, then, does the compatibilist get the idea that autonomy is a condition of responsibility? Not from the objective freedom of will, something that is ultimately unknowable. Furthermore, Kant would agree, even if no rational agent has ever acted freely, they would still be responsible for their actions, because in each case they could have acted in this way. How does one know they could? Certainly not by virtue of the agent's freedom, because such an argument would result in a *petitio principii*. Therefore, if the condition for assigning responsibility is the known possibility of free will, it could not be satisfied for the finite agent; only an omniscient agent could make this assignment. Thus, we see that the condition's strength is exaggerated, and Kant would not be disposed to accept all its consequences. Especially, he would not accept the consequence of interdicting all the assignments

of responsibility that we actually perform, and worse, of suspending the validity of all our moral judgments in the name of knowledge that human understanding cannot reach.

The negative pathway of morality, evil, provides the basis for another argument against the thesis that responsibility is not imputable save on the assumption of freedom of action. As regards human beings, Kant defines evil as a negative defect (*malum defectus*) and not as a positive evil of privation (*malum privationis*).<sup>1</sup> In other words, he defines evil as the *choice* of the maxim that satisfies the appetites, and not as a positive incentive to prefer a maxim of evil.<sup>2</sup> If this is so, then evil is the result of the non-intervention of good will—assumed, of course, to be a free will—since any choice requires some will. Consequently, with respect to the human creature, there is no autonomous action that has evil as a positive determination.

Now, if the assignment of responsibility implies the assumption that free will inaugurates causal chains, and according to Kantian philosophy this can only occur through respect for the moral law, then the human agent would only be responsible for good actions. Actions motivated by inclinations would contain in them, the inclinations, a positive incentive to choose interested maxims—in other words, maxims that are evil (*malum privationis*). Actions whose source of determination were not respect for the law would be heteronomous actions; one would not consider the agent responsible for them, since they would be caused by sensible motives.

The result of this argumentation is obscene for morals. Responsibility weighs on good actions, motivated by respect for the moral law, not on evil ones, motivated by natural law. Besides being obscene, the result is unacceptable from a Kantian perspective. Of course Kant wants to make us responsible for not choosing the preservation of autonomous will. And, of course, he wants to make us responsible for our bad choices.

The Kantian analysis of lies confirms this interpretation. The agent who lies, despite suffering empirical influences, is guilty of the choice he has made (A554–5/B582–3, my italics):

Ob man nun gleich die Handlung dadurch bestimmt zu sein glaubt: so *tadelt* man nichts destoweniger den Täter, und zwar nicht wegen seines unglück-

---

1 Cf. Howard Caygill, *A Kant Dictionary* (Oxford: Blackwell, 1995), 182.

2 Cf. the explanation Kant gives in *Religion within the Bounds of Bare Reason*, especially in Section III of the first part (6:27).

lichen Naturells, nicht wegen der auf ihn einfließenden Umstände, ja so gar nicht wegen seines vorhergeführten Lebenswandels, dennn man *setzt voraus*, man könne es gänzlich bei Seite setze, wie dieser beschaffen gewesen, und die verflossenen Reihe von Bedingungen als ungeschehen, diese Tat aber als gänzlich unbedingt in Ansehung des vorigen Zustandes ansehen, *als ob* der Täter damit eine Reihe von Folgen ganz von selbst anhebe. Dieser Tadel gründet sich auf ein Gesetz der Vernunft, wobei man diese als eine Ursache ansieht, welche das Verhalten des Menschen, unangesehen aller genannten empirischen Bedingungen, anders habe bestimmen können und *sollen*. Und zwar sieht man die Kausalität der Vernunft nicht etwa bloß wie Konkurrenz, sonder an sich selbst als vollständig an, wenn gleich sie sinnlichen Triebfedern gar nicht dafür, sondern wohl gar dawider wären; die Handlung wird seinen *intelligibelen Charakter* beigemessen, er hat jetzt, in dem Augneblicke, da er lügt, gänzlich Schuld; mithin war die Vernunft, unerachtet aller empirischen Bedingungen der Tat, völlig frei, und ihrer Unterlassung ist diese gänzlich beizumessen.

The agent is admonished, even if there are natural circumstances weighing on him that could explain the action from the point of view of natural causality. As one can see, freedom is indeed *assumed* by Kant, but not due to the argument that without it responsibility would vanish. Now, if not in this way, how does freedom enter the system? If freedom is from the empirical standpoint a *fata morgana*, what supports the assumption that this idea is effective?

### 3. Theory of Value

The conclusion derived from the argumentation above obliges us to ask about the origin of the idea—Kantian, doubtlessly—that rational agents can be free and autonomous. The touchstone for answering this question, in my opinion, lies in the theory of moral value that Kant strictly connects to his theory of duty. In *Groundwork* he teaches (4:1): “Nothing can possibly be conceived in the world, or even out of it, which can be called good, without qualification, except a good will.” Value, in turn, only begins with duty, and increases insofar as it—duty—is the reason for our action (4:11):

... würde er [der Menschenfreund] den nicht noch in sich einen Quell finden, sich selbst einen weit höhern Wert zu geben, als der eines gutartigen Temperaments sein mag? Allerdings! Gerade da hebt der Wert des Characters an, der moralisch und ohne alle Vergleichung der höchste ist, nämlich daß er wohlthut, nicht aus Neigund, sonder aus Pflicht.



That this is so does not result from some specious reflection, but is already clear in the moral judgments made by common people. Also in Kant, the concept of a good will “[is a notion] which exists already in the sound natural understanding, requiring rather to be cleared up than to be taught” (4:8).

The order of this argument proceeds in a manner contrary to that put forward by those who make responsibility depend on freedom. The assignment of responsibility is now what tells us that there must be another dimension of causality—a dimension enabling one to assume the will of the agent, because the person *should* have acted in a given way and did not do so. Where does the idea come from, that a person has the duty to act in a certain way? Again, it does not come from the idea that the person is free, since from an agent’s will alone, no free action can be recognized. Freedom is implied only by what is imputed to the action by all those who seek intelligibility in it. The judgment of value is what enounces the supposed existence of a free will and imposes the assignment of autonomy to the agent.

The Kantian theory of value, in turn, is supported by morality as practiced by common people. Considering the moral judgments that are ordinarily enounced, what distinguishes the value in them is duty, and even closeness to duty. In fact, the closer the action comes to the ideal of being motivated by mere duty, the more valuable it is; and, *mutatis mutandis*, the more autonomous. Admonishing the liar who, due to empirical circumstances, had reason not to be truthful only makes sense if one thinks about it against the background of a world where things should take place differently, an intelligible world, not a sensible world. The facticity of moral judgments is the manifestation of a reason that conceives the world beyond the constraints of natural laws—a reason that discerns an order that is not sensitive to things, and that, in this measure, conceives a law for them. This law is not for the order of nature, but for reason itself, insofar as it cohabits the natural world and insofar as it considers the practical element of its existence.

Daß diese Vernunft nun Kausalität habe, wenigstens wir uns eine dergleichen an ihr vorstellen, ist aus den *Imperativen* klar, welche wir in allem Praktischen den ausübenden Kräften als Regeln aufgeben. Das *Sollen* drückt eine Art Von Notwendigkeit und Verknüpfung mit Gründen aus, die in der ganzen Natur sonst nicht vorkommt. Der Verstand kann von dieser nun erkennen, *was da ist*, oder gewesen ist, oder sein wird. Es ist unmöglich, daß etwas darin *Anders sein soll*, als es in Allen diesen Zeitverhältnissen in der Tat ist, ja das Sollen, wenn man bloß den Lauf der Natur vor Augen hat, hat ganz und gar keine

Bedeutung. Wir könnten gar nicht fragen: was in der Natur geschehen soll; eben so wenig, als: was für Eigenschaften ein Zirkel haben soll, sondern was darin geschieht, oder welche Eigenschaften der letztere hat (A547/B575, my italics).

Practical law is an imperative for beings whose will is rational. It is respect to duty, required by moral admonishment, that requires the assumption of freedom and autonomy of will. The base of the metaphysical construction of morality, therefore, lies in the moral distinctions we all make, distinctions our judgments provide eloquent testimony of. Good, not freedom, reveals the kingdom of ends to us. Indeed, in Kant the good action, the good action in itself, requires freedom in the strong and metaphysical sense. Value, the absolute value, the good in itself, presents to Kantian theory the assumption of a supernatural will. The assignment of responsibility requires a weaker condition, the mere capacity to make moral distinctions, to recognize moral good (i. e., to recognize good as motivation by duty). Concerning the person, as Kant taught in the second *Critique*, “die Gewißheit einer Gesinnung, die mit diesem Gesetze [the moral law] übereinstimmt, die erste Bedingung alles Werts der Person ist” (5:130). Essential to personhood is, therefore, the disposition to act upon duty, as the very base for measuring a person's value.

#### 4. Conclusion

The fact of assigning responsibility is what places before us the need to presuppose that freedom is possible. The assignment of value corresponds to responsibility. To establish value, it is enough *for us to be able* to act freely; it is not necessary that this is the usual way men act. Indeed, in Kant one cannot find anything in the empirical world that is not determinately caused. Through experiences, then, we will not find anything, any action that is the effect of freedom,<sup>3</sup> which is beyond

---

3 “Wenn wir aber eben dieselben Handlung in Beziehung auf die Vernunft erwägen, und zwar nicht die spekulative, um jene ihrem Ursprunge nach zu erklären, sondern ganz allein, so fern Vernunft die Ursache ist, sie selbst zu erzeugen; mit einem Worte, vergleichen wir sie mit dieser in praktischer Absicht, so finden wir eine ganz andere Regel und Ordnung, als die Natur Ordnung ist. Denn da sollte vielleicht alles das nicht geschehen sein, was doch nach dem Naturlaufe geschehen ist, um nach seinen empirischen Gründen unausbleiblich geschehen mußte. Bisweilen aber finden wir, oder glauben wenigstens zu finde,

experience. Everything that freedom “causes” will belong to an intelligible world, not to the world of phenomena. This is the world of value. What requires the existence of freedom is, therefore, value, and this is attested by the moral judgments we indeed enounce, judgments whereby we impute blame to rational agents, when, being determined by natural appetites, they do not fulfill their duties.

An interesting issue, beyond the scope of this essay, would be to understand the empirical fact itself that is at the base of Kant’s assumption of freedom (*viz.* our moral judgments). The true challenge of a moral theory is not, as I see it, to make freedom and need compatible: in the absolute sense, these terms are contradictory, and in the relative sense, they are philosophically irrelevant. The real challenge is to explain the fact that a moral system exists, that we make mutual demands on each other, as though we were free, and judge each other as though our actions were not all part of a causal and natural system.<sup>4</sup>

---

daß die Ideen der Vernunft wirklich Kausalität in Ansehung der Handlung des Menschen, als Erscheinung, bewiesen haben, und daß die darum geschehen sind, nicht weil sie durch empirische Ursachen, nein, sondern weil sie dadurch Gründe der Vernunft bestimmt waren.” (A550/B578)

4 For further discussion on the nature of the system of morality, cf. A. N. de Brito, “Da validade de juízos morais: uma abordagem empirista”, in L. Guimarães, (ed.) *Ensaio sobre Hume* (Belo Horizonte: UFMG, 2005); and A. N. de Brito, “The Role of Reason and Sentiments in Tugendhat’s Moral Phylosophy”, *Crítica, Revista Hispanoamericana de Filosofia* 40 (2008), 29–43.

## 16. Moral Individuality and Moral Subjectivity in Leibniz, Crusius, and Kant

*Courtney David Fugate*

My aim in this essay is to explore the relation between what I will term moral subjectivity and moral individuality in Kant and two of his German predecessors. After defining and illustrating the distinction I make between individuality and subjectivity, as well as mentioning the associated complexities that arise from its use for interpreting common moral experience, I will argue two points: firstly, that while individuality and subjectivity, as I define them, were self-evident features of the moral agent in pre-Kantian German thought, they proved to be impossible to integrate into any traditionally conceived metaphysics of human nature; and secondly, that Kant's own Critical metaphysics of freedom is, at least in part, specifically designed to meet this challenge. The purpose of introducing this distinction and illustrating it by reference to this particular historical background is to show how a very similar difficulty to the one we find in the Reinhold-Sidgwick dilemma was already an issue even before Kant.<sup>1</sup> This will provide us with yet another way to begin seeing how Kant's theory is able to provide a solution. By way of conclusion, I will suggest that while Kant's Critical method succeeds in rendering individuality and subjectivity compatible, it still fails to provide any clear principle capable of justifying their unity. It is one thing to show that they can be joined into a single consistent image of the moral self; it is another still to prove that they must be so joined.

What then is the distinction between moral individuality and moral subjectivity? In its simplest form, I take individuality to be what distinguishes one thing from another. Just to cite one relevant source, Alexander Baumgarten in his *Metaphysica* glosses the individual as

---

1 The literature on this debate is extensive. See in particular: C. L. Reinhold, *Briefe über die Kantische Philosophie*, 2 vols. (Leipzig: G.J. Göschen, 1792), esp. vol. 1, 267–8; Henry Sidgwick, “The Kantian Conception of Free Will”, *Mind* 13 (1888), 405–12; Jens Timmerman, *Kant's “Groundwork of the Metaphysics of Morals”* (Cambridge: Cambridge University Press, 2007), 164–7.

being singular, or in other words, as completely determined.<sup>2</sup> He further says numerical difference, or the principle of individuation for things within a single species, is just the complete determination of those predicates that are left undetermined by the species itself. In accordance with this usage, I will initially define moral individuality in particular to be the unique moral character of a given agent (i. e., what specifically sets this agent apart from any other when the actions of both are characterized in strictly moral terms). Subjectivity, on the other hand, I take to be that inner active principle that first constitutes any event as an action or property, or makes it an action of a given thing that is consequently regarded as the *subject* of said action or property. Accordingly, I nominally define moral subjectivity to be that inner active principle that a moral agent employs to produce actions, or that enables such an agent to have specifically morally qualified acts imputed to it.

A sense of the difference I have in mind can be gained by extending Thomas Nagel's observation on the difference between the subjective and the objective views.<sup>3</sup> Moral individuality is composed, as it were, of every detail or objective moral fact about an agent (i. e., it requires complete determination), and from such details alone it is impossible to see any reason why I am this individual rather than another. This is because to say I am this individual rather than another does not add to the objective moral facts about that individual. Yet it remains meaningful to say I am some particular individual; indeed it is difficult to see what could be more meaningful. To say "John Smith is me" says nothing about John Smith, but it says much about me. Moreover, "me" and "John Smith" are not interchangeable *salve veritatem*, since "John Smith is John Smith" has an essentially different meaning from "John Smith is me"; this can be seen from the fact that the former is a tautology, while the latter is actually false. The significance of the statement lies not in an objective fact, but rather in how I reflectively situate my consciousness within and interpret this objective sphere.

A slightly more interesting formulation, also related to this distinction, will prove helpful in what follows. Firstly, I call *moral individualism* any theory that claims we can read off the inner nature of subjectivity—the very determining principle of its action—from a complete characterization of its individuality. I call *moral subjectivism*, on the other

---

2 See Baumgarten's *Metaphysica*, §148.

3 See Nagel's *The View From Nowhere* (Oxford: Oxford University Press, 1986), esp. ch. 3.

hand, any theory that holds the inner nature of moral subjectivity to be the only referent that enables one to judge or determine the individual moral character of an agent.

Furthermore, both moral individualism and moral subjectivism are basically teleological accounts of the structure of a moral agent. *Moral individualism* rests on a teleological inference from the pattern of the effects of the causality of a subject, to the character of this subject's governing purpose. *Moral subjectivism*, by contrast, rests on a teleological deduction, as it were, that begins with the governing purpose of an agent, and only from this deduces what effects are possible for an agent who takes this purpose as the law of its activity.

Now, in most non-moral cases these two teleologies are generally thought to hang together in a fashion that is quite trivial. If I stamp my foot every time I say a particular word, you will likely conclude from this pattern of behavior that I intended to do so. If, however, I tell you that I had no conscious intention of stamping my foot, then (unless of course you are Freud) you would likely revise your judgment and call it a tick, or a reflex—not an action but a simple physiological event. Thus, if I stamped on someone else's toe by the same mechanism, I could not be blamed. In such cases as these, we assume trivially that the intention is the purpose that causes the act, and the act is the effect that is a sign of the purposive character of the cause. The relation from causes to effects is deductive; the inference from effects to cause is imperfectly inductive; so in absence of the former, the latter gives way.

Say, however, that I repeatedly steal from my neighbor, but protest that I never intended to do so; I just found myself in a bad situation, where my mind for whatever physiological or psychological reason became clouded, and I suddenly found his property in my hands. What is interesting here, is that not only would others not accept this explanation, but I probably would not accept it myself (at least Kant thinks that morally we should not accept it, and in any case our conscience will not allow us to). This remains true even in the case that I am telling the truth when I say I never explicitly make the conscious decision to steal, and often in cases where I actually tried repeatedly to act on the opposite intention. Now, because we still cannot make sense of any act as being *my* act, without also thinking that an internal determining principle within my subjectivity has generated this act as mine, in these strange but certainly not uncommon cases we resort to speaking of hidden or even unconscious intentions. One would say that I, the unwilling thief, should stop deceiving myself, and take a look in the

mirror. *I should judge who I really am by reference to what I do*; that is, I should give up my protests for a principle of moral subjectivism and adopt the contrary principle of moral individualism. The inference to unconscious intentions here might well have an inductive empirical foundation, but the conclusion to my responsibility for this intention obviously stems from moral considerations alone. In any case, both inferences are clearly not based upon direct evidence; no one has ever been aware of or discovered an unconscious intention. This is just a hypothetical entity that we introduce to render our moral practice consistent.

Let us say, further, that I do just this, and following out the logic of moral individualism, I come to see that I am really a thief, an evil person indeed. I am now convinced that the moral character of my subjectivity is truly defined independently of my conscious intentions, and even in the most important cases my true character is often the very opposite of my firmest *conscious* intentions (or at least to what I silently say to myself as I steal). Now, since from this point of view the inner determining ground that I employ to produce an act as mine is something only accidentally, if at all, related to my self-consciousness, say that for this reason I further decide—*as would seem to be perfectly rational*—to give up trying to form any particularly strong intentions at all, whether for or against stealing my neighbor's things. Moral individualism in its pure form, it would seem, leads me to place who I really am outside of my own hands, as it were; my self-consciousness is just along for the unfortunately bumpy ride.

This too, however, is clearly an unacceptable outcome, morally speaking. Morality allows us just as little to retreat into the blissful position of being only who we most centrally intend to be (or tell ourselves we are), as it allows us to give up these intentions and admit that we are just whatever we happen to be. In neither case, it is clear, do we take proper responsibility for our actions, nor do we commit ourselves to becoming better persons in the future. In other words, explanation of the structure of the moral agent by reference exclusively to either one of these teleological accounts fails to capture the fundamentally more complex teleological structure that is essential to moral self-consciousness. The same effect has different explanations in each case. Moral individualism has this effect because it disassociates our self-consciousness from the moral quality of our behavior by treating our behavior itself as the independent principle of our characters. Moral subjectivism, on the other hand, also has the effect of undermining responsibility, but this

time because it defines our character by reference to our conscious intentions, and to this extent allows us to disclaim as really ours all those apparent behaviors others see us as committing, but that we from a deeper moral perspective do not see ourselves as really wanting to commit. Proper moral self-consciousness, I would contend, consists essentially in the purposive state of mind that results from the perceived need to resolve the conflict between the results of these two teleological accounts.

Now, the difficulty of constructing an accurate account of this tension within the confines of traditional metaphysics can be illustrated through a brief look at the theories of two of Kant's predecessors: Leibniz and Crusius. The foundation of Leibniz's account of the moral agent, of course, resides in his monadology. This is because the core of the moral agent is its soul, and this is nothing in Leibniz's view but a simple substance to which God has added apperception and reason. Like all simple substances, the soul contains within itself from eternity all that it will ever do, and the unity of the soul, the actual law of its activity, is nothing more than the productive unity of the *whole series of its actual determinations*.

One of Leibniz's chief concerns in the *Theodicy* is to show how God's foreknowledge is compatible with moral responsibility. Leibniz admits, of course, that we must be free in order to be responsible, so the issue resolves itself into the question: In what sense can a monad be said to be free, when like everything else, it is governed by the principle of sufficient reason? Leibniz's solution lies in the claim that moral agents are free precisely because the complete sufficient reason of anything that should transpire within their monads, including their immoral choices, lies entirely within these monads' own essences. The issue of freedom, as Leibniz explains, turns not on the necessity of an act occurring, but on the source of that necessity. If the non-performance of the act were to imply a contradiction, then it would be absolutely necessary and not morally significant. In all other cases, the ground lies in the essence of the agent itself, and so is properly a free act of that agent.<sup>4</sup>

In this way Leibniz's monadology, in conjunction with his concern to explain moral responsibility, leads him to what is perhaps the most

---

4 For instance, see G. W. Leibniz, *Theodicy*, tr. E. M. Huggard (Chicago: Open Court, 1985), 143–8, 151–5, 302–3. Note that I am speaking here only of the freedom required for responsibility, not Leibniz's other conception of freedom as self-mastery.



extreme form of moral individualism. For, according to Leibniz's metaphysics, the law of the inner activity of the agent is and can be nothing more than the inner representation of the complete series of its actual performances. What is more, the acts that make up the agent's determinations, insofar as their reciprocal coherence and harmony constitutes its perfection, are in fact the reason this particular subject was selected for creation.

The results of this radical moral individualism can be seen from the fact that conscious or rational intentions take no part in Leibniz's discussion of responsibility, or for that matter in his characterization of the freedom required for responsibility. This would seem to lead to a sort of moral fatalism, because it determines the actual ground of the character of one's subjectivity from a standpoint that one might not recognize, or may not ever actually become conscious of. Leibniz himself admits that we do not know why we do the majority of things we do—as he says, men are empirics, thus no better than brutes, in three-quarters of their actions—and yet we are absolutely responsible for everything we do nevertheless.<sup>5</sup> What is worse still is that our performance of bad actions occurs precisely in those impassioned instances where we lack a clear and distinct representation of ourselves. We must avoid, in Leibniz's view, a certain sort of fatalism that denies our actions and conscious intentions are themselves within the series of causes, and so does not recognize that they have a part in determining the character of our subjectivity. If you think, as some do in Egypt, that there is no reason to wear your seat-belt because if you are going to die then you are going to die in any case, then you have made the basic mistake of failing to see that not wearing the belt is in part the cause of whatever outcome there might be. Recognizing this, however, only avoids one sort of fatalism. For, admitting that our conscious intentions and strivings play some role in determining our character does not transform them into the central determining ground of subjectivity, as moral subjectivism and in part morality itself requires. They are simply accounted among the infinitely many other determinations that make up our individuality, the "individuating detail" as Leibniz calls it, and along with these are seen to belong to the subject by virtue of a sufficient reason

---

5 G. W. Leibniz, "The Principles of Nature and Grace, Based on Reason", in *Philosophical Papers and Letters*, 2 vols., ed. and tr. Leroy E. Loemker (Chicago: University of Chicago Press, 1956), vol. 2, 1036. See also, in the same volume, "The Monadology", section 28, 1048.

that is quite distinct from them.<sup>6</sup> To ask why I am this moral individual, say why I am for instance John Smith the thief, in Leibniz's system, is therefore transformed into the question as to what causes this individual monad to ask at this particular moment such an objectively empty and irrational question as *Why is John Smith, John Smith?* To want anything else than what actually happens is tantamount to wanting to be a different person; Leibniz claims this is simply absurd.<sup>7</sup> If only we add to this the recognition that it is best for things to happen in whatever way they

---

6 Leibniz, "The Monadology", sections 36–41, 1050.

7 Leibniz, *Theodicy*, 195. So what about guilt, moral struggle, moral progress, obligation, and similar moral concepts that would seem to require a more complex view of the moral agent, a view that can make sense of an agent that wants to be different from what it is? Leibniz, of course, has something to say on these issues, but again, the result is not very satisfying. His account proceeds by superimposing another structure, one that does take consciousness to be its core, over the underlying monadology. From the standpoint of the moral agent viewed simply as a monad, there is of course no difference at all between acting and representing, nor are there any differences in the degree of freedom or spontaneity that produce an act. Yet every such representation, in a monad with consciousness and reason, can have various degrees of perfection ranging from the most obscure and confused to the most clear and distinct. Leibniz claims that the clarity and distinctness of our representations of the causes of the actions that occur within ourselves are in direct proportion to our activity in regard to them. So despite the fact that all actions of a moral agent, viewed as a monad, are absolutely spontaneous and active, they appear as passive to us, and so as caused by an external source, to the extent that they are obscure and confused. Furthermore, actions whose reasons we only obscurely and confusedly represent are what Leibniz characterizes as slavish and unfree, while those whose reasons are clearly and distinctly conceived express our self-mastery and freedom. So here, finally, consciousness, implied by the notions of clear and distinct representation, takes up a place at the center of our moral image of ourselves. Moral progress, struggle, and obligation are all explained by Leibniz, and more fully by Wolff and Baumgarten, as arising from the contrast between actions whose reasons are confused and obscure and actions whose reasons are clear and distinct. In particular, the idea of a moral law would seem to be absurd when the agent is viewed as a monad, since the unavoidable law of such an agent is nothing other than the series of its acts. Leibniz, however, defines the moral law as what would be natural for a good person, meaning for a most perfect and thus most active one. Moreover, this law has a constraining force on our passions, only in the sense that knowledge of it will bring clarity into the causes of our own actions, and so will increase our own activity and freedom, thereby diminishing our overall passivity.

do inevitably happen, then we have the proper Christian fatalism that Leibniz equates with a joyful and peaceful state of mind.<sup>8</sup>

In Leibniz, at least, his metaphysical commitment to individualism clearly forces him to give a role to intentions that is not in accord with the role we need to give to them in common moral practice.<sup>9</sup> Our conscious intentions might not be everything, and they might under one description be among the moral facts about an individual, but for us to devote ourselves to morality, or to feel obliged to work consciously toward becoming better persons, such intentions must be given an intrinsically more central role by being connected up with an inner principle of subjectivity that somehow governs how any fact about this individual gets counted as among its moral acts in the first place.

Philosophers in what is now called the Thomasian–Pietistic tradition, a tradition that had considerable influence on Kant’s early work, rejected this Leibnizian view because they felt it turned our moral destiny into a matter of luck, and undermined any genuine sense of duty or law. Crusius, whose works Kant studied carefully, argues in particular for a metaphysical conception of freedom that would make sense of the inner experience of the moral struggle as a real struggle between what the moral self happens to be at any given moment, and what it knows it is supposed to become.<sup>10</sup> In contrast with Leibniz, Crusius thinks freedom is most properly understood to be conscious self-determination, and an action of freedom is in fact one whereby we determine our actions, not through an eternal essence that we may never even become conscious of, but rather through the very ideas we are conscious of, and nothing besides. The fundamental faculty of freedom, then, is nothing else in Crusius’ view than the faculty for determining our actions from ideas alone, and thus without any further incentive or prior determining cause.<sup>11</sup> What Crusius is attempting here, I believe,

---

8 Leibniz, *Theodicy*, 53–8, 151–5.

9 They should not be viewed as individual properties or actions of the monad, sufficiently determined by still other properties, that along with these latter all together constitute the character of the moral agent.

10 The fundamental idea of the Leibnizian line of thought, I have claimed, is that it is metaphysically incoherent to want to become anything other than what we are—that is, anything other than what is inscribed in our own individuated essences even prior to our having been created.

11 See Christian August Crusius, *Anleitung vernünftig zu leben* (Leipzig: J. F. Gleditsch, 1751), sections 37–56. See also his *Ausführliche Abhandlung von dem re-*

is to develop a sense of freedom with a determining principle that is thoroughly *internal* to self-consciousness, rather than internal to some metaphysically constituted eternal essence. He is trying to put his finger on a sufficient expression of the transparency condition of genuine teleological activity.

Crusius is brought to this notion of freedom in order, as I said, to make sense of our actual inner experience of moral struggle and choice, and our actual experience of the sort of absolute duty or moral obligation we know the moral law to be invested with.<sup>12</sup> His conception of freedom, we can now see, makes it possible to understand the moral law as what Kant calls a categorical imperative, precisely because it invests us with the capacity to act from the thought of the law alone, or for its own sake. To put the matter most briefly, Crusius' main contribution to the discussion lies in his emphasis on the essential link he sees between absolutely spontaneous, but also conscious, self-determination and the very notion of a responsible agent that is also governed by a categorical law. He recognizes, in other words, that there can be no pure moral individualism, such as Leibniz offers, since without the idea of an absolutely self-determining or causally self-transparent subjectivity at its basis, the individual would not possess the essential freedom of subjectivity *from all determinations of its individuality*; yet this is precisely what is required for a genuine sense of responsibility and law.

Crusius' extreme subjectivism, however, has its own difficulties, stemming from the emptiness and isolation of this fundamental faculty of freedom. The most serious problem in Kant's view is that the link between freedom and the law becomes inherently contingent. That is to say, because such subjectivity has no content, the law cannot be seen to follow necessarily from any fact about it. The law, if there is one, must therefore be some content that is derived from a source external to freedom itself.<sup>13</sup> There are simply no resources within Crusius'

---

*chten Gebrauche unter der Einschränkung des sogenannten Satzes vom Zureichenden oder besser Determinirenden Grunde* (Leipzig: C. F. Krausen, 1744), esp. sections 7–9. Here Crusius specifically addresses the views of Leibniz and Wolff.

- 12 This conception of freedom is able to deal with these issues precisely because it invests the human will with a faculty for determining itself to act according to the moral law, without however acting on any motive that would be separate from or prior to our actual recognition that we are obligated to act according to it.
- 13 Interestingly, Crusius does recognize that the law must be discoverable from within the nature of the human being (i. e., in our natural conscience), and

concept of freedom to explain how freedom *qua* freedom could possibly be motivated to adopt any particular law, let alone the moral law. Crusius is clearly forced into this position because he reasons that if we allow any individual character to determine the faculty of freedom, then we—to that very extent—compromise its inner self-determination, and thereby also compromise what makes it genuinely moral in this respect.<sup>14</sup> Interestingly, this also implicitly undermines any attempt to explain responsibility and moral striving, because the absolute capacity of freedom is only directed to deal with the character of my individuality because doing so is contained in the content of the law itself.

In this connection, note how Kant's own moral theory centers precisely on the problem of the relation between a proper concept of absolute or transcendental freedom and the specific content or law that is suitable for being its intrinsic form. Indeed, whereas Crusius comes to his notion of moral freedom exclusively from within the cosmological problem of an absolutely uncaused or first causality, Kant's analysis in both the *Groundwork* and the *Critique of Practical Reason* begins with a resolution of the idea of an absolutely unlimited act of *willing* (i. e., an unlimited act of causality according to representations). In this way, Kant is able, by a careful procedure of peeling back the limitations from such a causality according to representations (in this case, maxims), to arrive at the single representation or maxim that is itself unlimited and so is suitable to an unlimited causality according to concepts. By precisely this process Kant arrives at the equivalence of a free will and a will governed by the categorical imperative, through the recognition that the categorical imperative is the only unlimited or unconditioned representation fit to govern a free will. In other words, the former is the only content or form suitable to the idea of the latter. In the *Critique of Practical Reason*, moreover, Kant in fact says the idea of freedom would never even occur to us in a theoretical respect, if it were not first raised by our awareness of the unconditional character of the moral law (5:29–30, 163). But this claim essentially amounts to the further claim that absolute causality can only first be discovered by human reason through the original unlimited content or form that is the law. Freedom, in other words, could never be thought if it were merely

---

thus that it does not necessarily have to be revealed to us by God. So much is required by his commitment to Enlightenment values and the philosophical secularism of his own tradition.

14 Crusius, *Anleitung*, sections 46, and 211–3.

formless, if it were bare potency for any form, and it in fact is revealed to reason in the thought of an absolutely unconditioned form. This shows particularly clearly how Kant aims to retain Crusius' sense of absolute freedom while showing both that this same absoluteness precludes its taking on an externally-given form and that its nature as will requires it nevertheless to have a form intrinsic to it.

My thesis is that it was through direct awareness of the difficulties in Leibniz's and Crusius' accounts that Kant came to see that the morally upright individual must hold the principles of moral individualism and moral subjectivism together in their seemingly most mutually-exclusive forms, and that the dynamic tension generated thereby is what constitutes a genuinely moral state of mind. I think it was in this way that he also came to recognize that any account constructed within traditional metaphysics would inevitably fall into either absolute moral individualism or absolute moral subjectivism, and in either case would be unable to avoid leveling certain key aspects of moral experience. The difficulties of Reinhold and Sidgwick, and of a host of others, therefore arise in part from a failure to see how Kant's philosophy aims to resolve precisely this key issue, and Fichte's sensitivity to it is in large part due to his deep engagement with a project of explicating a dynamic account of consciousness similar to Kant's.<sup>15</sup>

Kant's account of the moral agent, I would therefore argue, is specifically designed to avoid the difficulties of previous metaphysical accounts, by its casting of the seemingly incompatible positions of moral subjectivism and moral individualism as dynamic moments in a self-developing and internally complex teleology of moral self-consciousness. What Kant does, in effect, is to lay out the internal logic that propels moral self-consciousness *from within* to take up both positions at different moments, and to hold them in a certain tension and balance for the sake of generating a properly moral state of mind, a state of mind that is aware of the moral as both *law* and also somehow as an expression of the character of one's own freedom. This is made possible by the way Kant's Critical method allows him to relativize these successive standpoints on our moral agency (through development of the concepts of both theoretical and practical objective reality), met-

---

15 See Johann Gottlieb Fichte, "Review of Leonard Creuzer, *Skeptical Reflections on the Freedom of the Will, with Reference to the Latest Theories of the Same, with a Foreword by Prof. Schmid*", tr. Daniel Breazeale in *Philosophical Forum* 32 (2001), 289–96.

aphysical though they may be, to the immanent unfolding of moral consciousness itself. The complex and purposive whole that is called moral self-consciousness, and is really crystallized in the intrinsically purposive and complex consciousness of the moral “ought”, is precisely the unity of this dynamic process.

I will not pretend, however, that Kant’s account is without its own difficulties. For although I think he renders these positions compatible through his interpretation of the metaphysics of morality as the inner normative structure of moral self-consciousness, and indeed even shows how they are logically interconnected—both major achievements—I fail to see how he can provide a justification of the overall unity of this process that does not either reduce his principle of autonomy to a mere *as if*, or else reduce the unity of moral self-consciousness to a theoretical hypothesis. But this is the subject of another essay.

Before concluding, I would like to draw one key lesson: any account that attempts, as many do, to extract Kant’s view of the moral agent from the specific logical order he employs to explicate it, and thus tries to cast it in a single metaphysical account minus the dynamism inherent in this logic of self-consciousness, will inevitably run into numerous seeming contradictions in what Kant says. Indeed, the upshot, I believe, is that any such account that tries to be consistent will end up leveling what is most original in Kant’s thought, and as a result will make him out to be either an absolute moral individualist or an absolute moral subjectivist. But in either case it will be unable to make any real sense of Kant’s account of duty, responsibility and moral progress.

PART IV

Judgment as the Orientation  
of Personal Unity





## 17. Aesthetic Judgment and the Unity of Reason

*Ulrich Seeberg*

### 1. The Unity of Human Personhood

According to Kant the unity of human personhood is based upon the unity of reason. By reason, a subject refers to the comprehensive idea of an intelligible world. This idea, as the thought of an unconditional totality of reality, has to include not only all physical objects and events of the sensible world but also the subject of thought and action as being aware of the sensible world. The idea of an intelligible world is therefore constituted by a subject's self-reflection on the common ground of the sensible world and of the subject as being aware of the sensible world. Since this self-reflective thought of reason is neither monolithic nor intuitively evident, the subject has to explore and justify its variants and their relations in order to unfold and stabilize the unity of reason. Hence, the unity of human personhood is not a given fact but has to be affirmed and developed by the subject's reasoning about the idea of an intelligible world. Kant compares this process, exhibited in his *Critiques*, with the construction of a self-stabilizing vault (5:3).

Kant distinguishes three different modes of reference to the idea of an intelligible world. The first *Critique* claims that a subject's objective knowledge *a priori* about the world is restricted to singular objects amongst others in space and time, the realm of nature. Although the corresponding idea of an unconditional totality of these objects and of their logical structure guides the subject's experiencing of the sensible world heuristically, this remains an ontologically undetermined and therefore problematic idea of theoretical reason. In addition, it is not possible to decide objectively if and how the subject as experiencing the sensible world belongs to an intelligible world that includes both, the subject and the sensible world. Responsible and free moral acts of the subject in the sensible world, on the other hand, presuppose that the acting subject thinks of itself as part of an unconditional intelligible world of freedom. Since the idea of human freedom can neither be proved nor disproved by theoretical reason, the categorical imperative

demands that we behave as if we were free. Nevertheless, the second *Critique* claims this demand of practical reason has its own, non-objective reality. Moreover, with respect to the idea of an ultimate end of human life, practical reason determines the idea of the intelligible world by the personal characteristics of intellect, good will, and omnipotence. Therefore, although we would not understand what it means to be free if we were able to explain freedom and the personal characteristics of the intelligible world objectively, the unity of theoretical and practical reason appears to be problematic. The third *Critique*, finally, claims that the awareness of a purposive relation between nature and subject, namely the judgment of beauty, bridges the gap between nature and freedom, respectively between theoretical and practical reason. Beauty assures there is a common ground of nature and the subject as being aware of nature. This awareness, as Kant holds, allows us to determine the theoretically undeterminable idea of an intelligible world by the concepts of practical reason and therefore stabilizes the unity of reason and of human personhood.

With a particular focus on aesthetic judgment, this essay aims to outline at least some of the basic principles of this quite complex and differentiated conception of the unity of human personhood. The first part recapitulates Kant's explanation of the tension between theoretical and practical reason by focusing on the question how the subject can grasp itself as part of a sensible as well as an intelligible world. The second part tries to explain how the judgment of beauty bridges the gap between nature and freedom. With respect to the question of how Kant's philosophy can be linked to the tradition of Asian thought, the third part finally and briefly asks if and how Kant's explanation of the aesthetic judgment can be related to the practice of meditation in the Buddhist tradition.

## 2. The Tension between Theoretical and Practical Reason

Kant's first *Critique* claims that constitutive knowledge *a priori* about the world is restricted to singular objects amongst others in space and time. What can be known *a priori* is that any possible object of experience is conditional on specific spatial and temporal relations between different objects. Pure theoretical reason transcends the sphere of these singular

sensible objects by the idea of an unconditional totality of objects being logically structured as a continuous hierarchical system of species and genera (cf. A652–7/B680–6). Kant argues that it is necessary to think that the empirical diversity of the sensible world is logically structured in this way in order to explore its properties systematically. This heuristic idea of a logical structure of the sensible world, however, does not amount to saying that the sensible world as such possesses such a structure. This point gains more clarity with regard to Kant's argument for the restriction of objective knowledge *a priori* in the transcendental deduction of the categories.

According to this argument the possibility of objective knowledge *a priori* depends on the unity of the subject's apperception.<sup>1</sup> The unity of apperception, the thought that I am necessarily the same subject of all my thoughts in order to know that I think a specific content right now, is bound to the subject's intuition of a given sensible manifold in space and time that cannot be grasped simultaneously but belongs, nevertheless, to a singular entity, namely nature.<sup>2</sup> The unity of apperception and the unity of nature form the two poles of a correlation. Therefore, the awareness of empirically different perceptions of nature presupposes that these perceptions are unified by the subject's thought of an objective correlate of them. This objective correlate, Kant argues, is one and the same in all recognition (A109) and is constituted by a synthetic function *a priori* that can be specified by the categories. However, the process of synthesizing different given perceptions as such has an infinite number of cases of application and gives rise to infinitely many possible objects of experience that all share an identical form. The performance of this synthetic function *a priori* requires the spontaneity of the subject's intellect as well as the receptivity of the subject's sensitivity in perceiving a given sensible manifold. Objective knowledge *a priori* about the world is therefore restricted to a given manifold of spatial and temporal perceptions of a subject being unified by the subject's thought of an objective correlate of these perceptions. This implies that neither nature altogether nor the subject's apperception can be

---

1 Cf. Dieter Henrich, *Identität und Objektivität. Eine Untersuchung über Kants transzendente Deduktion* (Heidelberg: Winter, 1976). Ulrich Seeberg, *Ursprung, Umfang und Grenzen der Erkenntnis. Eine Untersuchung zu Kants transzendentaler Deduktion der Kategorien* (Hamburg: Europäische Verlagsanstalt, 2006).

2 Kant makes this point by claiming that there exist not only forms of intuitions *a priori*, namely space and time, but also formal intuitions of space and time, giving the unity of representations (B160–1).

grasped as an object of recognition. Thus, any objective knowledge *a priori* that transcended the sphere of temporal and spatial perceptions toward the idea of an unconditional totality would destroy the subject's identity, and so also the unity of apperception. The idea of an intelligible world remains an undeterminable idea that can only serve as a heuristic thought for empirical research.

Since the subject qua subject is not itself a determinable object in space and time this restriction of theoretical reason implies that the subject's questions about its own status in the world, what concerns its origin, its ending, and its meaning, can neither be answered nor rejected by theoretical reason. The same is the case with analogous questions about the origin, the ending, and the meaning of perceived reality *per se* (cf. Avii). It may seem as if this diagnosis leads to agnosticism. We cannot justify objectively the idea that we as subjects belong to an intelligible world, including the sensible world (nature) and the subject as being aware of the sensible world. It may therefore even appear to be questionable if, or at least in what sense, a subject's thought of its own existence refers to a real entity at all. The decisive point of Kant's analysis, however, is that the restriction of objective knowledge *a priori* to a spatial and temporal manifold allows a critique of any attempt to identify the existence of recognizable objects in nature with the existence of reality *per se*. That our knowledge *a priori* about objects is limited means that reality includes more than we can recognize objectively. The idea of an intelligible world cannot be determined as objective knowledge *a priori*, but if there were no such idea we would not even understand what it means that our real knowledge *a priori* about the world is limited in principle. However, since we clearly know that we are not and cannot be simultaneously aware of all possible objects of experience, and so also that our intellect is not intuitive but discursive (cf. A722–5/B 750–3), we have to admit that our intellectual capacities are limited. Kant's argument quite convincingly shows that this limitation is not contingent and preliminary but constitutive for our self-understanding as subjects in the world. Although the reality of the subject qua subject seems to fade away if it is regarded as an object, it cannot be disputed that the subject's apperception implies a self-reflective thought of reason concerning its status as part of a comprehensive intelligible world.

This restriction of theoretical reason now opens the field for the claims of practical reason. Acting in a free and responsible manner means to start an empirically unconditioned new causal chain in the sensible world (4:446). Our moral deeds have empirical consequences but

they are not solely determined by empirical conditions. If this were not the case, we would have to grasp ourselves as mere objects in the sensible world. This, however, would destroy not only the unity of apperception of the cognizing subject but also the possibility of freedom and responsibility of the acting subject. The moral subject therefore thinks of itself as part of an intelligible world that includes the sensible world (cf. 5:103–6 and 4:452–3). Moreover, to justify one's acts means one regards oneself as member of a community of free and equal persons. Therefore, the categorical imperative demands that we do not treat other people only as means but also as ends in themselves and that we behave as if we were free members of an intelligible realm of ends in themselves (4:433–4). That this demand is not fictitious cannot be demonstrated by theoretical reason but, as Kant holds, moral obligation is a fact of practical reason in its own right (5:31–2). Furthermore, if we were able to explain the possibility of freedom, and so also the possibility of starting a new causal chain in the sensible world, we would not be obliged but would be saints who necessarily acted according to the idea of a realm of ends in themselves (5:93–4). The fact of being morally obliged by the categorical imperative on the one hand, and the restriction of our theoretical knowledge *a priori* about the world on the other hand, therefore correspond to each other. Human personhood does not consist of two separated segments but forms a complex whole that is unified by the relations between the different ways we reflect on ourselves as members of an intelligible world.

The exploration and stabilization of these relations, however, appears to be quite a difficult task. Kant argues that the moral dimension of human life also and necessarily entails the idea of its ultimate end (5:107–10). This idea not only implies the factum of moral obligation that has to be fulfilled but also the aim for happiness. Since happiness implies the feeling of the agreeable, and this means we experience ourselves as part of the sensible world, the relation between the intelligible world and the sensible world has to be specified further (cf. 5:110–3). The difficulty is that the ideal of happiness combines sensible inclinations (the feeling of the agreeable) and virtue (moral consciousness), yet these seem to contradict each other. The feeling of the agreeable implies that we experience ourselves as beings who are passively opposed to empirical objects that cause the feeling of pleasure in us. Feeling in this way means one does not understand oneself as part of an intelligible world but as the factual isolated center of experiencing the world by one's senses, and so also as being part of the sensible world only. The

feeling of the agreeable as such cannot be shared with other persons; on the contrary, they are potential rivals in the struggle for objects causing pleasure. This partly sensual nature of a person explains why morality appears in the form of the categorical imperative, demanding that one does not treat other persons according to one's sensible and therefore selfish inclinations. The aim for a responsible life and the desire for the agreeable do not automatically coincide. In addition, if one asks for the ultimate end of human life, pleasure as such cannot provide a satisfactory answer (5:442). This is so because pleasure is a factual and unstable experience that contradicts the idea that we, as beings who are able to live our lives in a free and responsible manner, asking for the ultimate end of a self-determined life, are part of an intelligible world. What then does Kant mean he claims happiness is combined with virtue?

Kant's well known but difficult answer is that virtue is the restricting condition of happiness, so that happiness cannot be thought of independently from virtue (5:114–9). Happiness is an ideal of contentment; as such, it cannot be anticipated or imagined as a concrete state of the empirical life. Happiness somehow implies sensual pleasure, although this cannot be imagined as a complete fulfillment of our sensible desires and wishes. The ideal of happiness, however, cannot be realized by our own means. All we can and have to do is to act morally, such that we can hope to be dignified in achieving happiness. The decisive point here is that the very ideal of happiness as such depends on morality. Morality, however, only leads to the dignity of being happy. Thus, morality alone does not represent a complete answer to the question of the ultimate end of one's life, we also aim to be happy. This, however, means we expect an intelligible world to be open and accommodating for us, as sensible and moral beings, being in need for happiness, and not only a sensible world that exists just factually and is potentially threatening for the isolated subject. Kant claims this expectation is linked with the idea of a supreme being. Firstly, the dignity to be happy has somehow to be judged or to be considered by a supreme or divine intellect; secondly, happiness cannot be distributed arbitrarily, but in correspondence to our idea of being happy with dignity, and this requires the good will of a supreme being; and, thirdly, we hope to achieve happiness by the support of a supreme being. Thus, the idea of an intelligible world is determined by the personal attributes of intellect, good will, and omnipotence (cf. 5:124–32).

To be sure, these personal attributes cannot be illustrated empirically; they represent an idea of practical reason in the mode of a reasonable hope and faith. Thus, one cannot simply object that these reflections are based on psychological needs and the augmentation of our empirical character instead of true insight. Kant stresses the point that the truth of this idea as a reasonable faith depends on the factum of moral obligation alone and that this factum can neither be proven nor disproven by theoretical reason. However, since the distinction between the sensible and the intelligible world does not mean there are two separated worlds, one necessarily has to ask how both of these aspects of a unique world can be reconciled. The empirical reality of the sensible world, including our own feelings of pleasure and pain on the one hand, is indisputable; but this reality appears to be just factual, providing no answer to the question about its ultimate end. The thought of an intelligible reality, as a matter of hope and faith on the other hand, provides an answer to that question; but this thought depends on our moral obligation, whose reality cannot be justified with reference to the empirical world. Although the restriction of theoretical reason and the demand of practical reason correspond to each other, there remains a tension between the realm of nature and the realm of freedom, and so also between the subject's awareness of being part of a sensible world and its thought or hope of being also part of an intelligible world.

### 3. Aesthetic Judgment as a Bridge between Nature and Freedom

The tension between nature and freedom is made explicit by the reflection of the faculty of judgment. Since the factual diversity of the empirical world seems to transcend any capacity of rational explication, the faculty of judgment has to presuppose that, nevertheless, the empirical world is somehow open or purposive for cognition. The idea of a logically structured empirical reality, however, has just a heuristic function for empirical research. The feeling of beauty is what reveals that there is indeed a purposive relation between self and world, so that a common intelligible ground of all reality is determinable by the concepts of practical reason (5:195–6). To clarify this function of aesthetic judgment Kant begins with an analysis of the difference between the agreeable and the beautiful. The agreeable is grounded on a contingent relation



between singular objects in space and time, on the one hand, and the sensible subject being passively opposed to these physical objects, on the other hand. If something causes the feeling of the agreeable, the subject is interested to maintain the existence of the causing object (5:205–7). This, however, is not a purposive relation between the subject qua subject and the sensible world but just a factual and contingent relation between physical objects in the world and the subject's feeling of pleasure or pain. The feeling of the beautiful in nature, on the other hand, does not imply any interest in the existence of something that could cause pleasant feelings in the subject. There simply is no physical object standing in causal relation to the subject's feeling, although there is a sensible world being perceived by the subject (5:204–5). In addition, beauty does not depend on a conceptual cognition of objects but on a subjective feeling that, contrary to the feeling of the agreeable, is experienced as commonly communicable. Thus, the meaning of the proposition "this rose is beautiful" is not logical or objective, but rather, pragmatically presents an invitation to other people to share the feeling of the beautiful, consisting in the awareness of a subjectively purposive relation between the subject qua subject and the sensible world (5:211–5).

Kant's explanation of these peculiarities of the feeling of beauty is that they are based on reflection of the faculty of judgment about its own relation to the sensible world. To make a judgment means to subsume special cases under a general rule. For this task, no further rule can be applied—the question how to apply a rule would simply occur again. How to make judgments cannot be taught; it requires practice and sure instinct for special situations (A133/B172). If one has to find new rules or concepts in order to judge a special case, the problem arises that there can be potentially infinitely many different concepts: there is no reason empirical reality in its diversity should correspond to the few concepts of human understanding (5:183). If empirical concepts and given intuitions coincided necessarily, one could not even distinguish between the contingency and the necessity of judgments. But clearly, empirical judgments are not necessarily true (5:403).

Given this situation, the contingency and the expected lawfulness of the empirical diversity can only be reconciled, as Kant argues, if one thinks the lawfulness of nature is somehow intentionally produced. To think that the idea of a specific causal or lawful relation precedes its reality means to think of an end or a purpose. Thus, the faculty of judgment reflects about nature as if nature or its intelligible ground pos-

essed intellect and will, and the principle of this reflection is that of the purposiveness of nature for the faculty of judgment (5:406–7). However, Kant emphasizes that there is no way to explain objectively this purposiveness of nature. What can be said is only that the attempt to apply concepts to nature presupposes, subjectively, thinking that these concepts and their systematic relations are applicable to empirical reality and that this seems somehow to work. Nevertheless, the task of judging each situation with sure instinct or intuition, the need to correct and refine our concepts, shows that we also have a sense for a non-factual and subjectively purposive harmony between our faculty of judgment and the situations where we apply these concepts. To judge appropriately, in accordance with specific situations, means we have to be attentive to the harmony between our faculty of judgment and the world. If this harmony is lacking, one feels that concepts are applied arbitrarily, dogmatically, or in an authoritarian manner. If one feels one cannot apply any concept at all, the world appears to be chaotic and meaningless, as when one dreams in a fever. This, presumably, is what Kant points to when he speaks of a task of the faculty of reflective judgment that is needed for every empirical cognition and also underlies the possibility of aesthetic judgment.<sup>3</sup>

It may seem as if this reference to the feeling of a harmony between self and world is not very illuminating. However, Kant offers an analysis of this feeling that allows us to recognize its philosophical importance, especially for what concerns the feeling of beauty. The feeling of a harmony between self and world is not just a factual subjective or private state caused by physical objects, but the result of an actively performed reflection of the judging subject about its relation to the sensible world. The awareness of this relation is a genuinely aesthetic phenomenon that, as Kant suggests, is generally based on the subject's reflection about the purposiveness of its representations for the preservation of its own state. Pleasure, Kant claims, just is the awareness of the purposiveness of a subject's representational contents for the subject's aim of self-preservation

---

3 Cf. 5:286–91. This argument has to justify the claim of the aesthetic judgment to be universally valid. However, my interpretation is that one can be aware of this function of one of the components of the aesthetic judgment as a feeling of harmony between a specific judgment and the situation being judged. Kant only explains that the aesthetic judgment refers to a purposiveness of a given intuition for cognition in general (cf. footnote 5, below). If this were a more detailed interpretation I would have had to show how to connect the notion of "cognition in general" with a specific empirical judgment.

(5:220). In the case of the agreeable, this reflection refers to factually existing physical objects in space and time and their causal impact on the subject's senses. But additionally, there is an accompanying reflection of the subject about these physical relations between the sensible world and the subject as being aware of the sensible world. This reflection brings to our attention that relations between subjective feelings of the agreeable and physical objects, as such, are contingent and unstable and therefore not purposive for the subject qua subject.<sup>4</sup> The same is the case with the application of empirical concepts. The faculty of judgment being aware of its capacity of applying concepts to the sensible world knows that the empirical diversity of reality transcends its capacity in principle. In the case of empirical recognition, being focused on specific objects in the sensible world, this leads to the heuristic idea of a logical structure of reality. The fact, however, that reality in its totality cannot be grasped or even intuitively imagined as an object of theoretical knowledge would just remain as an abyss for human understanding (cf. B641) if we did not have the idea of an existing intelligible reality as a guarantor of a purposive or meaningful relation between the sensible world and the subject who is aware of the sensible world. The reality of this idea, though transcending any theoretical or objective explanation, is already implied by the factum of moral obligation. But aesthetic judgment now confirms this idea as being constitutive for the feeling of the beautiful in nature.

The beauty of nature, as Kant explains, is based on the subject's reflection about the relation between itself as judging subject and some of the perceived formal qualities of the sensible world. If these perceived formal qualities (to be distinguished from the causal impact of physical objects on the subject's senses being felt as agreeable or not agreeable) are purposive for cognition in general, the result is a free and disinterested pleasure, a contemplation of beauty that can be shared with any other feeling subject.<sup>5</sup> Cognition in general means that the subject, by

---

4 Kant himself does not explicitly make this point. However, the agreeable as such would not present a problem for ethics if it did not imply the feeling of being factually isolated and conditioned. This feeling of being isolated and the need to harmonize sensible feelings with the intelligible world of freedom clearly is the result of a reflection of the subject about its status in the world and not a physically caused feeling.

5 Cf. 5:286–91. The notion of a purposiveness for cognition in general is quite controversially discussed; cf. Christel Fricke, *Kants Theorie des reinen Geschmacksurteils* (Berlin: de Gruyter, 1990); Hannah Ginsborg, *The Role of Taste in Kant's*

means of the faculty of judgment, is aware of its own capacity of cognizing or judging: to subsume intuitions as cases under concepts as rules, without any further rule guiding this activity of subsuming. To be aware of the purposiveness of the perceived formal qualities of intuitions for the general task of the faculty of judgment means the faculty of judgment actively but not arbitrarily judges the relation between given intuitions, as being perceived by itself, to its own capacity of subsuming intuitions under concepts. In this reflection the faculty of judgment provides the content, namely the perceiving of specific formal qualities of intuitions, as well as the principle of its judgment, namely the purposiveness of nature for the subject.<sup>6</sup>

To judge the purposiveness of perceiving specific formal qualities of intuitions for the general task of the faculty of judgment presupposes thinking there is an intelligible ground that establishes a harmonious relation between the subject as subject and the perceived reality as its purpose. Unlike the regulative function of the idea of a logically structured empirical reality in the first *Critique*, being purposive for particular objective judgments, the third *Critique's* idea of a purposiveness of perceiving specific formal qualities of intuitions for the general task of the faculty of judgment is constitutive for the subjective but universal feeling of the beautiful in nature. Thus, the feeling of beauty, linked with the idea of a meaningful intelligible world, including the subject as being aware of the sensible world, bridges the gap between the undetermined idea of an intelligible world of theoretical reason and the determined idea of an intelligible world of practical reason (5:195–6).

The feeling of beauty, however, presupposes that the sensible world as such stands in tension to the subject qua subject reflecting about its relation to the sensible world. If this tension simply disappeared, beauty would lose its meaning. Thus, the beauty of nature links nature and freedom, and by functioning in this way it also stabilizes the unity of human personhood. The identity of apperception, the factum of moral obligation, and the contemplation of beauty altogether form a complex that can and has to be explored and justified by reason.

---

*Theory of Cognition* (New York, London: Garland, 1990); Jens Kulenkampff, *Kants Logik des ästhetischen Urteils* (Frankfurt / M.: Klostermann, 1994).

6 Cf. 5:288. Cf. Hannah Ginsborg, “Interesseloses Wohlgefallen und Allgemeinheit ohne Begriffe (§§ 1–9)”, in Otfried Höffe (ed.), *Immanuel Kant. Kritik der Urteilskraft* (Klassiker Auslegen Bd. 33) (Berlin: Akademie Verlag, 2008), 4–77.

#### 4. Aesthetic Judgment and Meditation

Given the importance of the contemplation of beauty for Kant's explanation of the unity of human personhood, a very brief comparison with the religious practice of meditation seems to be interesting in order to explore the relation between Kant and Eastern thought. According to Buddhist tradition, the person who meditates has to transform all desires and wishes, memories and expectations, in order to experience the pure presence of reality. This presence and its meaning cannot be determined by empirical concepts but it can generally be characterized as a system of mutual relations between all parts of reality, such that no part can be conceived of independently from the others. There is no isolated object and there is no isolated subject; therefore no independent self exists.<sup>7</sup> This, interestingly, is quite similar to Kant's analysis of the beautiful. The beautiful in nature forms a complex that cannot be split into different parts and cannot be determined by empirical concepts; the contemplating subject of this experience is not interested in the existence of specific physical objects, and as being part of what it contemplates as beautiful the subject is aware of a common intelligible ground of all reality. Contemplating the beautiful in nature seems to represent a mode of meditation.

One might object that the notion of a subject does not fit very well into the Buddhist context. This certainly is a crucial point for any further discussion. However, it is revealing that the ethical ideal in the Western as in the Buddhist tradition demands not acting in a selfish way, and this entails overcoming the idea of an isolated self treating other persons just as means for its own purposes. Thinking of oneself as part of a common reality, acting in a responsible way, contemplating the beautiful—all this, and the deficient forms of it, can certainly be experienced by any human being, independently of any specific cultural background. The notion of a subject, therefore, only negatively refers to the idea of an isolated self or an independently existing substance. The positive content of this notion, particularly in the context of Kant's philosophy, cannot be determined objectively but requires one to think independently of oneself as being related to an intelligible ground of all reality. This, though it seems not to be remote from Buddhist teaching, leads to another difficulty: the determination of the idea

---

7 This point is stressed particularly by Thich Nhat Hanh, *The Heart of the Buddha's Teaching* (Berkeley, Ca.: Broadway Books, 1998).

of an intelligible ground of reality with personal qualities. Kant's thoughts are certainly influenced by the Christian religion, emphasizing that happiness cannot be achieved by means of a subject alone. But again, the idea of a supreme being does not refer to a substance that, in a rather questionable way, would be determined by empirical concepts. The idea of a supreme being is linked with the expectation that, in the end, sensible reality and the subject are not definitely separated, although the subject qua subject knows it is not identical with singular objects in the sensible world. Again, this does not seem to be incompatible with Buddhist thinking. Thus, the notion of a subject and the notion of a supreme being should not be understood as if they separated the traditions of Western and Eastern thought in principle. Instead, both points open a space for further discussion—and beauty may offer a comparatively unproblematic starting point for this undertaking.

# 18. Thinking with Instruments: The Example of Kant's Compass<sup>1</sup>

Nils Röllner

## 1. Introduction

During various phases of his life,<sup>2</sup> and even when he was 70 years old, Kant pondered theories of magnetism.<sup>3</sup> In Adickes' opinion, "Magnetism and electricity are precisely the areas where he gives his imagination free rein."<sup>4</sup> Adickes edited Kant's papers of scientific notes for the Berlin Academy edition, and he also wrote a two-volume monograph with the title *Kant als Naturforscher* (*Kant as a scientist*). Free to fantasize in the area of physics? This stands in diametric opposition to Kant's efforts to achieve orientation in metaphysical concerns. A characteristic of Kant's entire intellectual development is that he undertakes great efforts to achieve orientation. His life's work was devoted to the question: "What does it mean to orient oneself in thinking?" Ernst Cassirer sees this as Kant's basic question.<sup>5</sup> How does orientation in the *mundus intelligibilis* relate to the instruments of orientation that are used in the *mundus sensibilis*. This is the general question of the following discussion, and leads to the specific question: Is the compass an instrument that Kant "fanta-

- 
- 1 This essay is an abbreviated version of a text published in *Variantology 3—On Deep Time Relations of Arts, Sciences and Technologies in China and Elsewhere*, ed. Siegfried Zielinski and Eckhard Fülus (Cologne: Walther König, 2008). This essay, including quotes from German texts, is translated by Gloria Currence.
  - 2 Immanuel Kant, *Handschriftlicher Nachlass*, vol. XIV, nos. 23, 25, 26, 28, 29 (Berlin: de Gruyter, 1925), 80, 90, 94 f, 99 f, 106 f.
  - 3 Immanuel Kant, "Something on the influence of the moon on the weather condition" [1794], tr. John Richardson, in *Four Neglected Essays by Immanuel Kant*, ed. Stephen Palmquist (Hong Kong: Philopsychy Press, 1994); available online at [www.hkbu.edu.hk/~ppp/fne/essay4.html](http://www.hkbu.edu.hk/~ppp/fne/essay4.html).
  - 4 Erich Adickes, *Kant als Naturforscher II* (Berlin: de Gruyter, 1924–1925), 115.
  - 5 Ernst Cassirer, *Kants Leben und Lehre* [1918] (Darmstadt: Wissenschaftliche Buchgesellschaft, 1994), 42 f.

sized” about, or does he use the compass in pursuance of his philosophical search for orientation?<sup>6</sup>

## 2. Kant’s Scientific Notes

Kant’s scientific notes show that he was acquainted with the local declination and inclination of magnetic needles. He thought, for example, that the Earth was surrounded by a “universal magnetic atmosphere”.<sup>7</sup> He calls this the “sphere of heterogeneous ether”.<sup>8</sup> This sphere causes magnetic changes and the declination and inclination of magnetic needles.

Kant explains the polarity of a magnet as an arrangement of particles that have absorbed the magnetic atmosphere in differing concentrations.<sup>9</sup> These particles are attracted in varying degrees toward the centre of the Earth. Kant is trying to think through magnetism in a way that conforms to Newton’s laws. His notes on these speculations induce Adickes to label Kant a “cosmic wanderer” and speculative architect.<sup>10</sup>

In comparison to his scientist contemporaries Edmond Halley, Johan Carl Wilcke, or Johann Albrecht Euler, Kant is thus written off by Adickes as a dreamer. Kant ponders and speculates instead of poring over details of data and collecting facts exhaustively. He may have regarded the moon as a magnet, considered the movements of mountain ranges, and used all this to explain historical climate change, but his method disqualifies him as a scientist. Therefore, Kant’s notes on natural science can only be regarded as the speculations of a natural philosopher. Adickes bases his judgment on an analysis of notes that Kant did not intend to be published. Michael Friedman’s analysis is different. He works out the significance of the theory of ether for Kant’s later work.<sup>11</sup> The concept of ether is of central importance for Kant’s deliberations on magnetism. That contemporary theory of the magnetosphere of celestial objects, including Earth, does vindicate some aspects of Kant’s ideas on magnetism I shall not expand on here. For our discussion of the compass

---

6 Cf. Jordan Howard Sobel, “Kant’s Compass”, *Erkenntnis* 46 (1997), 365–92.

7 Kant, *Nachlass*, no.26, 96.

8 Kant, *Nachlass*, no.28, 99.

9 Kant, *Nachlass*, no.25, 90.

10 Adickes, *Kant*, 116.

11 Michael Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992), 290 f.



it is sufficient that Kant knew about magnetic phenomena, such as polarity, declination, and inclination, and that in this era a coherent explanation of these phenomena was lacking.

In the passages in his publications where Kant mentions the compass he does not discuss magnetic phenomena in detail or specific points about how compasses were manufactured. For example, he writes that winds “pass through the entire compass in fourteen days”.<sup>12</sup> In his scientific texts the compass is an instrument to aid spatial orientation, mentioned in addition to other means of orientation such as the stars. Kant was familiar with Francis Bacon’s notes on the magnetic needle and possibly also with William Gilbert’s *De Magnete*, criticized by Francis Bacon. Bacon accuses Gilbert of allowing experiments to infect his imagination.<sup>13</sup> On the “acus nauticae” Bacon writes in *Novum organum* (1620):

But we also note other discoveries of the kind which make us believe that mankind can pass by and step over outstanding discoveries even though they lie at our feet. For however much the discovery of artillery, silk thread, the mariner’s compass, sugar, paper, or the like may seem to stand on certain hidden properties of things and nature, there is surely nothing belonging to the printer’s trade which is not plain and pretty well obvious.<sup>14</sup>

That Kant was familiar with these thoughts is evidenced by a passage in *Anthropology from a Pragmatic Point of View*, where he considers the compass as an invention that humanity only needed to make once. Here invention is meant in contrast to discovery. Invention “brings into reality that which was not yet there” (7:247n), whereas through discovery something becomes perceptible “that was already there, for example, America, [or] the magnetic force directing to the poles.”

When the word compass is mentioned it occurs within the framework of what was common knowledge in Kant’s time and does not indicate extensive knowledge of this navigation instrument. Thus it is

12 Immanuel Kant, *Neue Anmerkungen zur Erläuterung der Theorie der Winde* [1756] (Berlin: Georg Reimer, 1910), 502; in this context see also Immanuel Kant, *Von dem ersten Grunde des Unterschiedes der Gegenden im Raume* [1768] (Berlin: Georg Reimer, 1912), 380. Cf. *Kant-Konkordanz zu den Werken Immanuel Kants*, ed. Andreas Roser et al. (Hildesheim: Olms-Weidmann, 1993).

13 Francis Bacon, *The New Organon* [1620], tr. Graham Rees with Maria Wakely (Oxford: Clarendon Press, 2004), I, no. 64, 101.

14 Bacon, *Organon*, I, nos.110, 160. Rees translates “acus nauticae” as “mariner’s compass”.

necessary to distinguish between Kant's preoccupations with the phenomenon of magnetism in his handwritten notes on scientific issues and the way he uses the compass in his published works on natural science. There, the instances when the compass is mentioned do not support the view that Kant is letting his imagination run away with him. What about the Critical writings?

### 3. Navigation Instrument of Critique

Thought is moved by the winds of doubt. A philosopher like Hume "deposited his ship on the beach (of skepticism)". By contrast Kant promises in *Prolegomena* (4:262) to give the ship "a pilot, who, provided with complete sea-charts and a compass, might safely navigate the ship wherever seems good to him, following sound principles of the helmsman's art drawn from a knowledge of the globe." In 1783 Kant speaks self-confidently of completeness. The region of metaphysics is charted and safe navigation of the ship is assured. Kant speaks as one who has brought to a conclusion ten years' work on the *Critique of Pure Reason* and is convinced that he has reliably mapped a new area of thought.

Soon, however, he no longer speaks of navigating charted waters. Navigation with compass and chart is possible in areas where other mariners have already been and recorded their positions. The domain of the metaphysics of morals is different (4:425): "Here, then, we see philosophy put in fact in a precarious position, which is to be firm even though there is nothing in heaven or on earth from which it depends or on which it is based." At first Kant develops the *Groundwork* of Critical ethics parallel to *Critique of Pure Reason*, as focuses on knowledge of nature. In *Groundwork of the Metaphysics of Morals*, published after the first *Critique*, Kant refers to a principle that, like a compass, facilitates orientation in moral questions. It is an aid, for example, in deciding whether a person should lie or not. A decision can be reached by using one's common sense. One only has to ask oneself if one would like to be lied to or not. When the question is put in this manner the person will make the right decision. The application of the principle, whether one can conceive of one's actions as a law for others or not, is like a compass: "Here it would be easy to show how common human reason, with this compass in hand, knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty or contrary to duty ..." (4:405).

The *Groundwork of the Metaphysics of Morals* clearly illustrates the difference between how reason was understood at the beginning of the Age of Enlightenment and the concept at the end of the eighteenth century. Reason can no longer refer conclusively—that is, according to objectively understandable rules—to a divine origin. At the same time it cannot refer to experience either. This means reason is suspended in a precarious position between earthly experience and heavenly guarantee. Reason has to guarantee its relevance itself. It cannot be derived from praxis or from predetermination.

#### 4. To Orient Oneself in Thinking?

What is the Critical method's attitude to religion? This is a question asked by Kant's contemporaries and enthusiastic readers. This question is also asked by the Prussian censor. Does Critical philosophy admit the conception of a reason that is greater than human reason? That is, does it admit a divine reason that is able to arrange rationally the world, humankind, and earthly events? When human reason thinks like this, then it reflects upon its own limits and thus also upon the limits of Kant's method.

Kant takes up a position on this by speaking of rational faith. This faith derives from "the right of reason's need, as a subjective ground for presupposing and assuming something which reason may not presume to know through objective grounds" (8:137). Knowledge through objective grounds is limited in the field of supersensible objects. Nevertheless, the need to orientate oneself is felt. The compass—and this is what is specific about Kant's compass—is an instrument that offers orientation in this field. As an instrument, it is made by humans. Yet it utilizes forces that in Kant's era were imperfectly understood. Kant equates such an instrument with rational faith (8:14):

A pure rational faith is therefore the signpost or compass by means of which the speculative thinker orients himself in his rational excursions into the field of supersensible objects, but a human being who has common but (morally) healthy reason can mark out his path, in both a theoretical and a practical respect, in a way which is fully in accord with the whole end of his vocation; and it is this rational faith which must also be taken as the ground of every other faith, and even of every revelation.

It is striking that Kant refers to navigation instruments when he is talking about the newness of the Critical method. In this way he gives his cri-

tique the status of a discovery, and endows his arguments with the aura of a radical change. His thought is thus elevated to the same level and import as the discovery of new continents. Do his references to the compass, therefore, signify that he has let his imagination run away with him? Could Kant have chosen some other means of orientation? In its forays into speculative territory can reason orient itself on the sun or the “polestar”? Is rational faith comparable to a “plumb line”, or do mainsprings act within reason? Could numbers have been used to offer orientation to reason in its precarious situation?

Kant is acquainted with various means of spatial orientation. In some of his central arguments he mentions counting. Counting, for example, promotes decisively the thought of subjective consciousness. Kant mentions the plumb line frequently in contexts where the potential of reason is expressed.<sup>15</sup> His references to the compass delineate an area where reason comes up against limits of objective grounds. Kant is so well-acquainted with the compass and magnetism that here he can use an equivalent for his argument. The compass is a product of human art. This art utilizes principles that are not fully understood. In this the compass is similar to reason: in the sense that it first has to develop in humans as a capacity for cognition, reason is a product of humans. At the same time reason has needs that are felt: for example, to assume a divine order. This felt need or rational faith is not compared with a sign post that someone has put up, but with a specific instrument, a compass, that one has to take in hand oneself. And the compasses in Kant’s day quivered a lot in the hand, as a visit to the Hellmann Collection today will confirm.<sup>16</sup>

The comparison between mainsprings and the compass is a central element in what is specific about Kant’s compass. Mainsprings are components of artifacts. They are a part of an organized whole. Kant thinks of nature as such a planned and organized whole. In *Groundwork* he calls it a machine (4:438). In nature one finds no superfluous elements. Kant

15 Cf. A762/B790, where Guyer and Wood translate “Richtschnur” with “guidance” (the German word connotes also the sense of measurement), and 5:16, where Gregor and Wood translate “Richtschnur” with “standard”.

16 Hans-Günther Körber, “Katalog der Hellmannschen Sammlung von Sonnenuhren und Kompassen des 16. bis 19. Jahrhunderts im Geomagnetischen Institut Potsdam”, in *Jahrbuch 1962 des Adolf-Schmidt-Observatoriums für Erdmagnetismus in Niemegeke* (Berlin: Akademie-Verlag, 1964), 149–71. I should like to thank Hans-Joachim Linthe for the tour and for information about the Hellmann Collection.

also mentions mainsprings when he speaks of subjective motives, when humans are thought of as machines and act like machines. The compass is not thought of with reference to the world machine. The laws that guide the functioning of the compass cannot be deduced as part of the overall plan of nature. Thus the compass is an instrument that points beyond the order of things that is understood. Kant thinks of the compass as an instrument that brings reliability and uncertainty together. Mainsprings, on the other hand, are a part of a strictly controlled system; their function is assigned to them by a superordinate whole that is conceived of as a machine. The compass functions as an aspect of an incomprehensibility that is nevertheless reliable.

Remarkably, in this realm of abstract arguments feelings and felt needs are spoken of. This is a zone of indeterminacy where mainsprings, as a component of mechanical arts, are out of place. In “What does it mean to orient oneself in thinking?” the compass leads to a place where cool reason becomes sympathetic—sympathetic in the sense that it guesses it is possibly dependent on determinations. Reason enters into an accord with these determinations. Naturally, this is not sympathetic for those, like Nietzsche, who philosophize with a hammer. To philosophize with a hammer instead of a compass means to want to hammer in order to receive information or—in the reception of Nietzsche in the *George Circle*—it is equivalent to hammering into the world the values one has set for oneself.<sup>17</sup> This is something entirely different to adjusting concepts of orientation when confronted with the limits of one’s own method.

Kant develops his concept of reason as a faculty humans utilize for self-determination. At the same time, by making use of the compass, he acknowledges that this faculty has its limits. This is the point where Kant acknowledges something that reason cannot comprehend and cannot objectivize. From here expands the area where Kant’s compass offers orientation. But Kant conceives the orientation with the compass as not per se sufficiently reliable. This can be deduced by his reflections on means of navigation, discussed in his *Physical Geography*.

Of the mariners of his time Kant demanded that they know the regions of the world at all times. That was the opinion of the “landlubber” thinker in his lectures on physical geography at the University of Königsberg. Seafarers can orient themselves “at any time” if they correct for the declination of a compass by regularly taking their bearings from

---

17 Ernst Kantorowicz, *Kaiser Friedrich der Zweite* (Berlin: Georg Bondi, 1927).

the stars (9:307; no.81). In conjunction with other reference systems the compass is useful. It is not the mediation of two systems that is interesting but rather the difference between two systems. In navigation this offers orientation.

In Kant's compass is inscribed a basic insight into the limited nature of human cognition, and at the same time an insight into the fruitfulness of thinking the differences. This latter is an insight that acknowledgment of the limitedness of human cognition is reconcilable with structural openness toward the other. Sharing this insight is a condition for communication and mediation between subject and object, between different subjects and their different cultures. This is an insight migrants and nomads regard as principal. That the "landlubber" thinker, "the great Chinaman of Königsberg",<sup>18</sup> promotes this insight is a challenging feature of his idealism. We may face this challenge by discussing the instruments that guide our orientation. Thinking instruments prepare us for the deep sea voyage that we all have embarked on already—especially the compass, a Chinese invention.<sup>19</sup>

---

18 Friedrich Nietzsche, *Beyond Good and Evil. Prelude to a Philosophy of the Future* [1886], tr. Helen Zimmern (New York: Dover Publications, 1997; online: <http://users.compaqnet.be/cn127103/ns/select.htm>), VI, no. 210: "Even the great Chinaman of Königsberg was only a great critic." On Nietzsche's relationship to Kant, see Stephen Palmquist, "How 'Chinese' was Kant?", *The Philosopher* 84 (1996), 3–9; available online at [www.hkbu.edu.hk/~ppp/srp/arts/HCWK.html](http://www.hkbu.edu.hk/~ppp/srp/arts/HCWK.html).

19 The history of the compass is discussed under the aspect of orientation in Nils Rölller, *Magnetismus—Eine Geschichte der Orientierung* (München: Fink, 2010).

## 19. Common Sense and Community in Kant's Theory of Taste

*Bart Vandenabeele*

### 1. Universal Communicability

Kant's theory of taste suggests the possibility of community with others who share our sensibilities and capacities to judge the beauty of nature and art. This community is, however, not based on common concepts, arguments, meanings, opinions, or convictions, nor on dialogue, contingent social and historical coherence, nor on general moral agreement or rational law. By considering the possibility of an aesthetic common sense (*sensus communis aestheticus*) that is fundamentally enmeshed in human *sensibility* and *affect*, Kant introduces the idea of an affective "sensus communalis" that not only aims to justify the universal communicability (*Mitteilbarkeit*) of aesthetic judgments, but also ultimately joins together the most personal contingent feelings of pleasure (or displeasure) with a necessary, universal, and cosmopolitan idea of humanity, based on the shareability of affects.<sup>1</sup> This transcendental idea has to be presupposed, or so Kant argues, in order to legitimate the communicability of aesthetic pleasure and promote human sensibility and affect in communion with others (5:297, 433, 355).

The problem of founding the aesthetic judgment's universal validity claim can only be tackled by introducing an important term that Kant uses for the first time in the notorious §9 (on the question whether in a judgment of taste the feeling of pleasure precedes the judging of the object or the judging precedes the pleasure), viz. universal *communicability* (*allgemeine Mitteilbarkeit*). This is what needs to be accounted for if

---

<sup>1</sup> Though I cannot argue here for this claim, I would suggest that the view put forth by Rachel Zuckert, that the conception of non-conceptually governed community is "a sphere of shared *meaning*—i. e., culture (in the non-Kantian meaning of the word)", is based on a misunderstanding of Kant's considered view concerning the social nature of taste. See R. Zuckert, *Kant on Beauty and Biology* (Cambridge: Cambridge University Press, 2007), 382 (*italics added*).

we want to find the justification for the universality claim in the judgment of taste. Unfortunately, Kant has written one of his most confusing passages on this very issue. He states (5:217):

If the pleasure in the given object came first, and our judgment of taste were to attribute only the pleasure's universal communicability to the presentation of the object, then this procedure would be self-contradictory. For that kind of pleasure would be none other than mere agreeableness in the sensation, so that by its very nature it could have only private validity, because it would depend directly on the presentation by which the object *is given*. Hence it must be the capacity for being universally communicated of the mental state [*allgemeine Mittheilungsfähigkeit des Gemüthszustandes*], in the given representation, which underlies the judgment of taste as its subjective condition, and the pleasure in the object must be its consequence.

Kant makes two rather puzzling statements. First, the pleasure is said to be the result of the aesthetic judgment; but how is this possible if the pleasure is also supposed to be the judgment's condition or ground? Secondly, aesthetic pleasure is argued to be the consequence of the universal communicability of the mental state in the judgment. How can the pleasure of taste be the consequence of the universal communicability of the mental state, when the latter is supposed to be pleasurable itself (at least in positive judgments of taste)? This looks really circular.<sup>2</sup> But, in fact, Kant can be defended by making an essential distinction between the act of judging or contemplating the object (*Beurtheilung des Gegenstandes*) and the judgment of taste (*Geschmacksurtheil*) as such. Judging the object obviously precedes the pleasure, but the latter precedes the actual judgment of taste. It not only precedes it, it also forms the determining ground of the judgment of taste proper. According to Paul Guyer,<sup>3</sup> §9 contains the basic elements for a theory of aesthetic appraisal that consists of two logically—but not necessarily phenomenologically—distinct acts of reflection: first, an act of mere reflection where pleasure is felt; and second, an act of aesthetic judgment proper where the pleasure is attributed to the harmonious play of the faculties. On this reading too, however, Kant's apparent implication that the universal communicability of the mental state in judging the object is itself

2 The following is partly based on D. Crawford, *Kant's Aesthetic Theory* (Madison: University of Wisconsin Press, 1974), 70, and H. E. Allison, *Kant's Theory of Taste* (Cambridge: Cambridge University Press, 2001), 111 f.

3 See P. Guyer, *Kant and the Claims of Taste* (Cambridge: Cambridge University Press, 1997), 8, 51–2, 97–102, 140–2, 147, 228.



the source of the sensed pleasure is, as Guyer puts it, “obviously absurd”,<sup>4</sup> since it suggests that universal communicability is constitutive of aesthetic pleasure instead of merely playing a part in evaluating it.<sup>5</sup>

Yet in the same section Kant explicitly denies that pleasure in the ability to communicate one’s mental state could account for the aesthetic pleasure in the beautiful. There is, Kant says, pleasure in the ability to communicate, but one cannot appeal to it to explain the transcendental necessity connected with the pure judgment of taste, since the former is merely an empirical or natural propensity (*Hang*) to sociability (5:218). To explain Kant’s remarkable implication, that aesthetic pleasure is grounded in universal communicability, we perhaps ought to qualify the aesthetic judgment in the way Hannah Ginsborg does, viz. as “a formal and self-referential judgment that claims, not the universal validity of an antecedently given feeling of pleasure, but rather its own universal validity with respect to the object”.<sup>6</sup> On Ginsborg’s account, a judgment of taste is a judgment *about* the normativity of one’s own mental state. The demand for assent is merely the demand that others recognize this normativity (i. e., that I judge the object as it ought to be judged, namely as beautiful). However, one might wonder how self-referential judgments could avoid making use of concepts, and hence, whether Ginsborg’s account does not illegitimately turn aesthetic judgments into intellectual judgments—that is, judgments of cognition requiring concepts to determine the correctness of ascribing one’s mental state

---

4 Guyer, *Kant and the Claims of Taste*, 137.

5 As Guyer contends, this would imply that in a solipsistic situation no one could take pleasure in a beautiful object. Only if there were the possibility of communication would aesthetic pleasure be possible. This was actually Kant’s anthropological view before he wrote the *Critique of Judgment*. See *Logik Blomberg* (24:45–6): “taste can therefore impossibly be separately solitary [*abgesondert eigenthümlich*]”; see also *Logik Philippi* (24:353–5) and *Anthropologie Collins* (15:179–80). This also occurs, however, in texts written after the *Critique of Judgment*, as in, for instance, his *Anthropology from a Pragmatic Point of View* (7:244) and the *Metaphysics of Morals* (6:212). See also *Metaphysik L<sub>1</sub>* (28:249–51), where he argues that the universal sense (*allgemeiner Sinn*) underlying judgments of taste has to be identified with a communal sense (*gemeinschaftlicher Sinn*) and also emphasises that “whoever does not come into a community has no communal sense” (28:249).

6 H. Ginsborg, “Reflective Judgement and Taste”, *Noûs*, 24 (1990), 72. See also her “On the Key to Kant’s Critique of Taste”, *Pacific Philosophical Quarterly*, 72 (1991), 290–313., and “Kant on the Subjectivity of Taste”, in Herman Parret (ed.), *Kant’s Aesthetics / Kants Ästhetik / L’esthétique de Kant* (Berlin: de Gruyter, 1998), 448–65.

to others, or, at least, into what Kant calls in his *Logik Dohna-Wundlacken* “beautiful cognition” (*schöne Erkenntnis*); such cognition is no longer based on the free play of the cognitive faculties, and is, therefore, altogether different from a pure judgment of taste.<sup>7</sup> Secondly, it is hard to see how Ginsborg’s view could allow for negative judgments of taste: if Kant meant the judgment of taste to be self-referential (i. e., *referring to the normativity or appropriateness of my mental state with regard to the object deemed beautiful*), and if the pleasure of taste is really *in* the universal communicability of my pleasure, then there is no room for a universally communicable *displeasure*, since, as Allison rightly notes, “universal communicability is itself a source of pleasure”.<sup>8</sup>

I do not believe Ginsborg’s view is what Kant had in mind. It is hard to see how one can account for the intricacies of Kant’s theory of aesthetic response without logically distinguishing between two acts of reflection. On the other hand, it is equally difficult to see how Guyer’s logical distinction can be translated into more phenomenological or “psychological” terms. Phenomenologically speaking, what Kant intimates seems to be the following: judging (*beurtheilen*) the object results in a feeling of pleasure or displeasure in me, and this feeling, transcendently grounded in the harmonious play of imagination and understanding, signals its universal communicability by means of its affective purity (i. e., its disinterested nature). Although it forms no legitimate basis for the universal validity or communicability of the judgment of taste proper, the disinterestedness is actually the affective “symptom” of the fact that the pleasure (or displeasure) must be attributed to the reciprocal quickening of the mental faculties that are operative in aesthetic judgments of taste—the “feeling of life”<sup>9</sup> of the subject—and not to some idiosyncratic inclination or quirk: displeasure signals the dishar-

---

7 See *Logik Dohna-Wundlacken* (24:710, emphasis added): “Wenn Anschauung und Begriff zusammenstimmen zur Belebung der Erkenntnis selbst, so machen sie in uns ein Wohlgefallen und dann nennt man es schöne Erkenntnis. Man muss sich bemühen, dass Verstand und Einbildungskraft zu einem Geschäft zusammenstimmen. *Dies ist aber nicht mehr Spiel.*”

8 Allison, *Kant’s Theory of Taste*, 115.

9 For an interesting treatment of the parallels and differences between beauty’s “feeling of life” (*Lebensgefühl*) and morality’s “feeling of spirit” (*Geistesgefühl*)—the latter is not a feeling of sense, although it is in some way palpable—see John H. Zammito, *The Genesis of Kant’s Critique of Judgment* (Chicago: The University of Chicago Press, 1992), 292–305. Interestingly, Kant also uses the term *Geistesgefühl* and not *Lebensgefühl* in connection with the feeling of the sublime.

mony, whereas pleasure signals the harmony of the two cognitive powers involved in aesthetic judging. It is in this sense that Kant's statement, quoted above, can be readily understood: "it must be the capacity for being universally communicated of the mental state [*allgemeine Mittheilungsfähigkeit des Gemüthszustandes*], in the given representation, which underlies the judgment of taste *as its subjective condition*, and the pleasure in the object must be its consequence" (italics added).

Instead of qualifying this idea as hopelessly absurd, as Paul Guyer does, or (like Hannah Ginsborg) defining aesthetic judgments as self-referential, thus depriving them of their disinterested nature, since pleasure *in* the universal communicability of the mental state cannot be disinterested but is (as Kant holds) a natural inclination, or, as Henry Allison<sup>10</sup> does, trying to solve the problems in the quoted passage by changing the words, as if Kant had committed a slip of the pen, one might perhaps interpret Kant's words in §9—no matter how clumsy Kant's formulation is—as follows. Phenomenologically speaking, the purity of taste, the disinterestedness of the experienced pleasure, is subjectively determined (*as its subjective condition*, Kant says) by the necessity of being universally communicable or shareable (i.e., by the "signal" in the mind [*Gemüth*] that the felt pleasure [or displeasure] is universally communicable). Thus, the universal communicability of the mental state is the affective sign of the purity of the felt pleasure or displeasure. This implies that, on the one hand, the disinterestedness of the pleasure is the essential, *a priori* condition for the universal communicability of aesthetic judgments: without disinterestedness aesthetic judgments could not be universally shared. But, on the other hand, the capacity for universal communicability is itself the ideal gauge to estimate whether the experienced pleasure is really disinterested or not. Hence, Kant writes: "the pleasure in the object must be its consequence", meaning that only in and through its possible universal communication can the felt pleasure be estimated to be actually disinterested or not. So in this sense, and in this sense only, can the pleasure in an object be the consequence of "the capacity for being universally communicated of the mental

---

10 Allison, *Kant's Theory of Taste*, 115: "In reality, however, all that is required to rectify matters is to substitute 'a universally communicable mental state' for 'the universal communicability of the mental state.' In addition to removing much of the air of paradox around the text, this would allow room for the possibility of negative judgments of taste; for there is nothing inherently problematic in a universally communicable mental state of displeasure (as opposed to a displeasure in its very communicability)."

state”. Whether or not the pleasure is really pure pleasure (i. e., the disinterested pleasure that grounds a pure judgment of taste) depends on the very universal communicability of the aesthetic judgment, the *ratio cognoscendi* of the disinterestedness of the pleasure. That one can actually be fairly certain (though one will never be able to prove it by means of arguments) that the pleasure one experiences here and now is disinterested is grounded in—though not caused by—the universal communicability of the mental state, or more precisely, in the affect that “signals” whether or not the activity of the mental powers is universally communicable. And this affect, or rather this universally communicable mental state, of course, presupposes “a capacity for being universally communicated”.

## 2. Aesthetic Common Sense

Since judgments of taste are not cognitive judgments, they cannot be based on a determinate objective principle, and hence they are not unconditionally necessary. On the other hand, if they had no normative force, “if they had no principle at all . . . , then the thought that they have a necessity would not occur to us at all” (5:238). This (subjective) principle is called “common sense” (*Gemeinsinn*) in §20. Kant stresses that this aesthetic *Gemeinsinn* must be clearly distinguished from the common understanding (*der gemeine Verstand*), for the latter judges according to concepts and not according to feeling. There is a crucial difference between the “sensus communis logicus” (mentioned in §40) and the “sensus communis aestheticus”.<sup>11</sup>

---

11 There is a striking parallel between the way Kant introduces the problem of the *sensus communis* and the central argument on the conditions of mathematical knowledge in his *Prolegomena*. (See Allison, *Kant's Theory of Taste*, 148–9.) There Kant is concerned with how mathematical knowledge can be both synthetic and *a priori*, and he will show that this is only possible under the assumption that an underlying pure or *a priori* intuition is possible. Kant is well aware of the paradoxicality of such a notion. He writes (4:282): “There is therefore only one way possible for my intuition to precede the actuality of the object and occur as an *a priori* cognition, namely if it contains nothing else except the form of sensibility, which in me as subject precedes all actual impressions through which I am affected by objects. For I can know *a priori* that the objects of the senses can be intuited only in accordance with this form of sensibility.” In the *Analytic of the Beautiful*, the problem of the universal shareability of taste is no less paradoxical: how can something intrinsically private as a *feeling* really claim universal

The aesthetic common sense he is talking about in the context of judgments of taste is first and foremost a *sense*, so it is a question of sensibility and feeling, “a sense (or feeling) for what is universally communicable, which can also be assumed to be universally shared. Otherwise expressed, it is a shared capacity to feel what may be universally shareable.”<sup>12</sup> This notion of common sense, although introduced rather abruptly in §20, has a long history in Kant’s own views on the communal nature of taste, as is clear from the following passage in the *Lectures on Metaphysics* (28:249):

But how can a human being pass a judgment according to the universal sense, since he still considers the object according to his private sense? The community among human beings constitutes a communal sense. Out of the intercourse among human beings a communal sense arises which is valid for everyone. Thus whoever does not come into a community has no communal sense.—The beautiful and the ugly can be distinguished by human beings only so far as they are in a community. Thus whomever something pleases according to a communal and universally valid sense, he has *taste*. Taste is therefore a faculty for judging through satisfaction or dissatisfaction, according to the communal and universally valid sense. But taste is still always only a judging through the relation of the senses, and on that account this faculty is a faculty of pleasure and displeasure. Objective satisfaction or dissatisfaction, or judging objects according to universally valid grounds of the power of cognition, is the higher faculty of pleasure and displeasure. This is the faculty for judging of an object whether it pleases or displeases from cognition of the understanding according to universally valid principles. If something is an object of intellectual satisfaction, then it is good; if it is an object of intellectual dissatisfaction, then it is evil.—Good is what must please everyone necessarily.—But the beautiful does not please everyone necessarily, rather the agreement of the judgment is contingent.

---

validity? As Henry Allison rightly contends, “the idea of a common sense, as the only condition under which such a claim regarding a mere feeling is possible, plays precisely the same role in the case of taste as that of pure intuition does in the case of mathematics” (Allison, *Kant’s Theory of Taste*, 149). This is as far as the analogy goes, however, since, as is well known, Kant did not invent the term *sensus communis*, whereas he did introduce the technical term “pure intuition”. Yet, *sensus communis* was used before Kant in a totally different way: in (especially) Scottish philosophy, common sense referred to common understanding, whereas Kant is here concerned with a *sensus communis aestheticus*, and he will not claim that taste can be based on a logical common sense, the *gemeiner Menschenverstand*. This is not surprising, since in the *Prolegomena*, he utterly disparages any appeal to this notion to solve the Humean problem of causality.

12 Allison, *Kant’s Theory of Taste*, 149.

This is a very interesting fragment for several reasons. Kant had clearly not developed his complex idea of the “exemplary necessity” of pure judgments of taste (i. e., the necessary requirement of such judgments’ universal assent), that is not based on concepts (5:237). But equally clear is that in his *Lectures on Metaphysics*, as in his *Anthropology*, the idea of the close relationship between the capacity of taste and being-in-community is already overly present, and that Kant had already developed the thought of an affinity between the notions “sensus communis” and “sensus communalis” before writing the *Critique of Judgment*. In the *Critique of Judgment*, however, Kant develops the idea of a “sensus communis” within a transcendental framework; this explains why he cannot ground the universal shareability of judgments of taste in the human inclination to sociability (*Geselligkeit*). On the other hand, it also warns against too strict a separation of those approaches: Kant would never have come up with the idea of a *sensus communis*, if he had not been convinced of the anthropological existence and relevance of togetherness and being-in-community: hence he says “the beautiful and the ugly can be distinguished by human beings only so far as they are in a community”; an idea that Hannah Arendt<sup>13</sup> will develop in her famous political reading of Kant’s third *Critique*. Moreover, Kant was well aware of the pleasure of being with others and of the interest of sociability. The whole idea of universal communicability (*Mittelbarkeit*) follows from this intuition: the peculiar knowledge that what I feel could be and ought to be shared by other people is itself a source of pleasure.<sup>14</sup>

We are ready now to follow the argument Kant develops in §21. The argument can be divided into seven different steps:<sup>15</sup>

---

13 H. Arendt, *Lectures on Kant’s Political Philosophy* (Chicago: University of Chicago Press, 1982).

14 It goes without saying that, as Béatrice Longuenesse argues, the fact “that this pleasure serves the empirical interest in sociability is a derived fact that no more diminishes the disinterested character of the pleasure, than does the fact that aesthetic pleasure generates an empirical interest in surrounding ourselves with beautiful objects.” See Longuenesse, “Kant’s Theory of Judgment, and Judgments of Taste: On Henry Allison’s *Kant’s Theory of Taste*”, *Inquiry* 46 (2003), 155.

15 I take these different steps literally from Allison, *Kant’s Theory of Taste*, 150–1. See also Lyotard, *Leçons sur l’Analytique du sublime* (Paris: Galilée, 1991), 242–4, for a similar analysis.

1. Cognitions and judgments, together with their accompanying convictions [propositional attitudes] must be universally communicable. This is a condition of claiming agreement with an object; consequently, its denial leads to skepticism.
2. This entails that the mental state required for cognition in general, that is, the “attunement” [*Stimmung*] of the cognitive faculties, which is that “proportion” [*Proportion*] suitable for turning representations into cognitions, must also be universally communicable. Again, to deny this would be to open the door for skepticism, since this attunement is the subjective cognition of cognition.
3. This attunement actually occurs whenever the perception of a given object puts the imagination into play, which, in turn, sets the understanding into action; but this attunement varies in proportion to differences in the occasioning objects.
4. Nevertheless, there must be one optimal attunement, that is, one in which the inner relation is most conducive to the mutual quickening of the cognitive faculties with a view to cognition in general; and this attunement can be determined (recognised) only by feeling (since the alternative—concepts—is ruled out).
5. Moreover, both this attunement and the feeling of it in connection with a given representation must likewise be universally communicable.
6. But the universal communicability of this feeling presupposes a common sense.
7. Consequently, we do have a basis for assuming a common sense, without relying on psychological observation, as a necessary condition of the universal communicability of our cognition, which must itself be presupposed if skepticism is to be avoided.

A lot could be said (and has been said) about this argument, especially if we read it as an attempt to provide a deduction of common sense as a condition of taste, but I just want to focus on one element that is crucial if we want to develop Kant’s idea of a common sense in a more pragmatic (and hence less speculative) way. It is important to note that Kant here claims merely to have shown that there are grounds for assuming a common sense, not that he has proved the *existence* of a common sense. He fully realized that he needed much more argumentation to provide a proper demonstration for that. One problem with his account in §21 is that it seems to suggest a solution he wanted to avoid, viz. that the aesthetic common sense must be presupposed as a condition of cognition in general. This Humean solution is surprising, since it is hard to grasp how a common sense that was defined in §21 as the effect arising from the *free play* of our cognitive capacities could serve as a condition of cognition. This is, as Allison argues, clearly incoherent.<sup>16</sup>

---

<sup>16</sup> See Allison, *Kant’s Theory of Taste*, 153.

There is only one way of avoiding this incoherent reasoning, and that is by assuming that in §21 Kant uses the term “common sense” not to refer to the effect of the free play of our cognitive powers, not to taste as such, but rather (as Allison rightly holds) to the capacity “for immediately seeing (without appeal to rules, and therefore through ‘feeling’) whether, and how fully, a given intuited manifold accords with a particular concept, that is, judgment”.<sup>17</sup> By analogy, it can then be argued that there are *epistemic*, non-psychological grounds to assume that in taste too there must be a *sensus communis aestheticus*; that is, one must first assume there is “a peculiar talent” to recognize a fit between understanding and imagination in cognition, otherwise the assumption of the capacity to do so when the cognitive powers are in free play (as is the case in taste) is at least highly implausible. Kant clearly has not sufficiently demonstrated that we do have a basis for assuming or postulating an aesthetic common sense, but indirectly he has provided arguments for its plausibility, by showing that “the attunement itself, and hence also the feeling of it, must be universally communicable” (5:239)—unless we plainly endorse skepticism, as Kant, of course, does not. The question now arises as to whether the link with morality might provide a better ground for the aesthetic judgment’s claim to universal assent.

### 3. Aesthetic and Moral Normativity

In §22 Kant argues that aesthetic judgments contain an “ought”: using common sense in matters of taste implies not that everybody *will* agree with my judgment but that they *ought to*; therefore, the demand for agreement in judgments of taste is comparable to similar moral or cognitive claims. Donald Crawford,<sup>18</sup> Ray K. Elliott,<sup>19</sup> and many others have argued that the demand of a judgment of taste for universal agreement is a moral claim, but although Kant clearly argues that there is a moral interest in beauty (and especially in natural beauty), the requirement that judgments of taste be universally shared is not of a moral but an *epistemological* kind.

---

17 Allison, *Kant’s Theory of Taste*, 154–5.

18 Crawford, *Kant’s Aesthetic Theory*, 143, 145, and *passim*.

19 Ray K. Elliott, “The Unity of Kant’s ‘Critique of Aesthetic Judgment’,” *British Journal of Aesthetics* 8 (1968), 244–59.



Nevertheless, we must distinguish two different “oughts” in taste: first, the epistemic demand for universal assent. As Allison says, “it is this ought that presupposes common sense, and it is quite independent of morality.”<sup>20</sup> The second “ought” is the demand to acquire and develop the capacity of taste itself, and this demand or requirement *is* connected with morality. But, *pace* Savile, it does not follow that the first “ought” is based on the second. Naturally, it would be absurd to demand universal assent to a judgment of taste, unless I presupposed that there are others with a similar ability to discriminate between beauty and mere charm or agreeableness. But from this it does not follow that the requirement to develop and refine these discriminatory abilities grounds the aesthetic “ought” (i. e., the demand for universal shareability of judgments of taste). On the contrary, and here I am again in agreement with Allison, the moral demand to develop the capacity of taste *presupposes* the universalisability claim of the pure judgment of taste: only because there is an inherent normative claim in judgments of taste could it be morally interesting to develop the capacity to make such judgments.<sup>21</sup>

This *caveat* is important, especially because Kant has made many commentators believe that the transition (*Übergang*) from beauty to morality is equivalent to saying that taste is inherently moral or that the capacity of taste is a necessary condition to a good will. Nothing of this is actually faithful to Kant’s enterprise, though. On the contrary, he clearly insists on the universalisability of taste, and even claims—against any humanistic interpretation of the *sensus communis*—that taste is more entitled to being called “common sense” than is the “common human understanding” that reduces the idea of communality to vulgarity, and whose possession involves no merit whatsoever (5:294). Instead, Kant makes clear that his talking of common sense refers to “the idea of a public sense [*gemeinschaftlichen Sinnes*]”; that is (5:293–4):

a power to judge that in reflecting takes into account (a priori) ... everyone else’s way of presenting [something], in order as it were to compare our own judgment ... not so much with the actual as rather with the merely possible judgments of others, and [thus] put ourselves in the position of everyone else, merely by abstracting from the limitations that may happen to attach our own judging.

---

20 Allison, *Kant’s Theory of Taste*, 159.

21 See Allison, *Kant’s Theory of Taste*, 159.

Although there are clear analogies here with moral reflection, Kant does not claim that abstracting from the limitations of our aesthetic judging (i. e., paying attention solely to the formal features of our representation or our representational state) is a necessary or sufficient condition for morality. The latter would lead to aestheticism and is obviously not what Kant wants to defend here. On the contrary, taste and morality are both autonomous; only because taste lays claim to universal assent may it symbolize morality. The adoption of a *general perspective*, a perspective that takes other people's judgments democratically into account, is crucial in this sense: in virtue of this capacity to adopt a broadened point of view, as it were, aesthetic reflection, bringing along "a certain ennoblement", may serve as a symbol for morality. As Henry Allison writes,

just as the beautiful does not effect a transition from the sensible to the supersensible because it symbolizes morality, but rather symbolizes morality because it effects such a transition, so, too, the pure judgment of taste does not make a valid demand on others because it symbolizes morality, but rather it is because of the "purity" underlying the validity of its demand that it symbolises morality.<sup>22</sup>

#### 4. Conclusion

In aesthetic judging we not only experience ourselves as free from the push and pull of our drives, needs and desires but also experience the "liberality" of our active engagement with an object and an enlivenment and expansion of our cognitive capacities (5:268). This experience of "aesthetic freedom", as grounded in the freedom of the imagination in aesthetic judging, is no merely private or personal matter. It is, Kant rightly holds, inextricably connected with the feeling of being-together with others, with whom we share similar capacities to judge the beauties of nature and art. In this sense, beauty presupposes not only a shared aesthetic sensibility but also a subject that is not primarily concerned with its own personal and private sensations but is always, at least in principle,

---

22 Allison, *Kant's Theory of Taste*, 267. Here I am in agreement not just with Allison, but also with Salim Kemal, "Aesthetic Necessity", *Kant-Studien*, 74 (1983), 184 and A. C. Genova, "Aesthetic Justification and Systematic Unity in Kant's *Third Critique*", in G. Funke & T. M. Seebohm, eds., *Proceedings of the Sixth International Kant Congress* (Washington, D. C.: Center for Advanced Research in Phenomenology & University Press of America, 1989), vol.2, pt.2, 293–309. This reading goes against Crawford, *Kant's Aesthetic Theory*.

aware of the shareability of its aesthetic affects and the communality of its aesthetic appraisal. In aesthetic judging, we do not merely feel our “elevation” above mere sense pleasure but also experience, as it were, a common affective bond with a cosmopolitan community that all human beings participate in and whereby they “esteem the value of others” (5:353). Aesthetic judging is, therefore, not of signal importance because it would make a transition from the sensible to the supersensible, but because it offers a way of rethinking subjectivity and intersubjectivity as manifest within felt, particularized pleasures that are universally shareable—not despite but *due to* their affective nature.<sup>23</sup>

---

23 5:353, 355. I wish to thank Stijn Van Impe for his invaluable comments.

## 20. Aesthetics and Morality in Kant and Confucius: A Second Step

*Christian Helmut Wenzel*

In the framework of his transcendental philosophy Kant strictly separates morality from aesthetics. The pleasure in the good and the pleasure in the beautiful are two different kinds of pleasure (*zwei Arten des Wohlgefallens*). As a consequence, a moral act as such cannot be beautiful. Only in a second step does Kant indicate possible connections between morality and aesthetics in his comments on aesthetic ideas, symbolism, the *sensus communis*, and education in general. In Confucius by contrast, we do not find such a radical separation between beauty and morality. He talks of humaneness (*ren*, 仁) and ritual (*li*, 禮). Projecting Kantian notions into the Analects, “beauty” seems to slide between the two and “moral” acts appear to be beautiful. One might wonder whether Confucius missed a point, or whether Kant overdid the separation. Or maybe both conceptions, of morality as well as of beauty, cannot so easily be translated from one philosophical tradition, or mind, to the other, and there is nothing like *ren* and *li* in Kant. In this essay I ask whether there is an “inner” and an “outer” in Confucius, and I introduce Kant’s notion of “subjective purposiveness” and relate it to the Confucian notions of *dao* (道) and *tian* (天) as well as to *ren* and *li*.

Reading the Analects, one easily feels that Confucius trusts in certain correlations between the inner and the outer, where I think here of the inner as moral feeling or *ren* (humaneness, benevolence) and of the outer as *li* (ritual).<sup>1</sup> We should practice (outer) rituals to acquire the right (inner, moral) attitudes. One feels this suggestion is based on the belief that outer performances can make us aware of inner feelings for human values. Of course, there is no guarantee: “The Master said, Clever words and a pleasing countenance—little humaneness [*ren*]

---

1 My association of the inner with moral feeling or *ren* is intended to be loose and preliminary. Differences and similarities should become apparent in the course of this essay.

there.”<sup>2</sup> Things can go wrong, and we all know that. Confucius certainly did. Nevertheless, some kind of trust in positive correlations can always be felt throughout the Analects: “The Master said, A human being who lacks humaneness—what is ritual to someone like that? A human being who lacks humaneness—what is music to someone like that?”<sup>3</sup> Music and ritual must be performed with the right attitude, because only through such an attitude do they acquire their true meaning and value. Confucius believes in outer practices leading to the development of such inner attitudes. The performance of music, archery, and other forms of ritual will lead one to realize, see, and develop the right kind of moral views and feelings. For this reason, and with this hope, Confucius recommends education and learning in general. The outer is not merely a sign but also a stimulus for development of the inner. This is what I mean by “trust in positive correlations”.

My talk of the “inner” and the “outer” may seem imported and projected here. One might think Confucius himself does not speak in such terms.<sup>4</sup> He does not speak of “mental representations” or “souls”. But he knows of deceit and mere outward appearance; that is all I need here to feel justified in importing these terms of an inner and an outer for the moment.<sup>5</sup>

Comparable, one might think, to the correlation of beautiful ritual and moral attitude in Confucius, Kant talks of a “beautiful soul” (*eine schöne Seele*) (5:300 [§42]) and of “beauty as a symbol of morality”

2 Confucius, *The Analects of Confucius*, tr. Burton Watson (New York: Columbia University Press, 2007), 1:3.

3 Confucius, 3:3.

4 But see 4:17 (“reflect on *your own* account”) and 5:27 (“anyone who can ... *look inside* himself, and put the blame there ...”). Confucius does talk about the inner, in some way. How, and to what degree, has to be worked out.

5 Hence I tend to side with Benjamin Schwarz against Herbert Fingarette. See Benjamin I. Schwarz, *The World of Thought in Ancient China* (Cambridge, Mass. and London: Harvard University Press, 1985), 72–4. For a passage pointing out sources in Herbert Fingarette, Benjamin Schwartz, Henry Rosemont, and Chad Hansen, discussing Confucius’ concern (or lack of concern) for people’s internal psychological life, see Philip J. Ivanhoe, *Ethics in the Confucian Tradition* (Indianapolis/Cambridge: Hackett Publishing Company, 2002), 170, n.22. Also Tu Wei-Ming, *Humanity and Self-Cultivation: Essays in Confucian Thought* (Boston: Cheng & Tsui Company, reprinted 1998; Lancaster-Miller Publishers, 1978) freely talks of an inner and an outer, of *ren* as an “inner morality”, a “self-perfecting” and “self-fulfilling process of an individual”, and a “principle of inwardness”, and of *li* as its “externalization” (9–13, see also 17–30).

(5:351 [§59]). Kant even provides a whole theory that allows us to explain how beauty and morality are linked and why and how this link is merely an indirect one, thus leaving, and even creating, room to account for failed correlations. According to Kant, beauty and morality are based on acts of reflection where we take ourselves as subjects in general, abstracting from personal considerations. Pleasure in beauty must be disinterested and felt by us as human beings as such. Only then are we justified in claiming universal validity for our judgment of taste. Similarly, the morally good is realized in acts based on universality considerations regarding rules and principles.<sup>6</sup> This similarity, or isomorphism, between aesthetic and moral acts of reflection can be imagined to be underlying some of Confucius' intuitions. In *Analects* 3:8, for example, he writes:

Zixia asked, saying,  
*Her artful smile engaging,*  
*Lovely eyes in clear outline,*  
*Colors on white ground,*

What do these lines mean?

The Master said, The painting comes after the white background.

Zixia said, So ritual comes afterward?

The master said, Shang (Zixia) is the one who reads my meaning. At last I have someone to discuss the *Odes* with.

I read this passage as a way of pointing out *ren*: three items are given and the fourth, *ren*, has to be figured out: As the lovely (*mei*, 美) eyes are related to the white ground, so is ritual to *ren*. The latter is in both cases a prerequisite for the former; *a* stands to *b* as does *c* to *x*, where *x* is the unknown element:  $a:b = c:x$ . Without the white background (*b*), you cannot draw the eyes in clear outline (*a*), and without *ren* (*x*), ritual (*c*) does not make sense. The eyes are beautiful only against the white background and ritual must be performed with the right attitude. It is the white background in relation to the colorful and lovely eyes that should remind the reader of *ren* as a prerequisite for ritual and its beauty. The analogy is an aesthetic one. Beauty serves morality by being a visible symbol for it. This way of putting it fits the Kantian conceptual framework of beauty being the symbol for morality.

---

6 See Christian Helmut Wenzel, *An Introduction to Kant's Aesthetics: Core Concepts and Problems* (Malden, Massachusetts and Oxford, England: Blackwell, 2005), 113–9, for an exposition of beauty and morality and their symbolic relationship according to Kant.

But Confucius does not talk about “acts of reflection” or a “free play of our faculties of cognition”. He is not a philosopher like Kant. Why and how morality is a prerequisite for beauty of ritual, such as archery or music, remains unclear. Nevertheless, the picture Confucius offers has its charm (not in Kant’s sense of *Reiz*, though usually translated as “charm”; 5:223 [§13]). The eyes, colors, clear outline, and white ground all belong to the realm of the visual and aesthetic, whereas ritual and *ren* belong to the moral realm. But there is no such strict separation here. Ritual itself is aesthetic, and the white ground is given a moral tone. Both ritual and *ren* appear as being beautiful. Kant would, strictly speaking, resist such a view, or move, and allow for it only in a symbolic way. Moral feelings cannot be perceived and therefore cannot be beautiful. But, on second thought, maybe *ren* is not quite the same as moral feeling.

Another connection between the moral and the aesthetic, if we continue to allow ourselves to import these terms here, can be seen in a passage relying on the notion of harmony (*he*, 和): “The Master said, What ritual values most is harmony. The Way of the former kings was truly admirable (*mei*) in this respect.”<sup>7</sup> But what exactly is meant by “harmony” here? The context talks of loyalty, trustworthiness, and filial behavior as being of primary concern, whereas ritual, the arts, and government come second.<sup>8</sup> I therefore think the harmony of inner attitudes is meant here and gives the ritual its beauty. Such beauty has to “shine through”. But how is this supposed to work? Rituals are performed in the form of acts and processes that unfold in time and space. Seen from the outside, one has to know, or somehow “see”, that they are not superficial and merely punctually performed but that they reflect and are the result of the right inner attitude. How does one “see” such an inner attitude from the outside? Even seen from the inside, in first-person perspective, when performing rituals oneself, one is led, as one might be surprised to find out, to the same question: One often imagines oneself as being seen and regarded by others. One sees oneself through the imagined eyes of others and thus relies on others to see oneself. The inner is accessible in first-person perspective partly through imagining a second- or third-person perspective. Of course one can go wrong, or even deceive oneself. There is room for vanity. But that is another question.

---

7 Confucius, 1:12.

8 See Confucius, 1:2, 1:4, 1:7, 1:11, and 1:13.

“Zixia asked about filial devotion. The Master said, The difficult part is the facial expression.”<sup>9</sup> Again, Confucius trusts in the correspondence. The harmony that “ritual values most” (quoted above) can therefore be understood as the outer appearance of an inner harmony, similar to the facial expression that should be the outer appearance of the proper inner attitude of filial devotion. Confucius’ trust in education through ritual is based on such inner-outer correspondences. Another passage that I read in this way is 4:1: “The Master said: Humaneness is the beauty of the community.” In opposition to the interpretation suggested so far, an interpretation that takes harmony and beauty as being based on inner qualities such as loyalty, trustworthiness, and filial behavior, one could quote 3:25. There Confucius talks of Wu music as being “perfect in beauty, but not perfect in goodness” and thereby presupposes the possibility of beauty being separated from inner qualities.<sup>10</sup> Apparently something can be beautiful without the right inner qualities shining through. But I think 1:12 and 4:1 are not meant in this way.<sup>11</sup> In fact, they point out the (ideally) right correspondence between beauty and goodness and the desirability of that correspondence.

Another example of a link, tacitly assumed between the inner and the outer, can be found in 6:16: “The Master said, If you have the good looks (*mei*) of Song Zhao but lack the eloquence of Invocator Tuo, you’ll have a hard time escaping blame in the world today.” Confucius lived in difficult times, when having good looks and even being a good person were often not enough. Although it is not something that Confucius approves of, he recognizes it as an unfortunate fact that sometimes you have to argue and be eloquent if you want to be successful and escape blame “in the world today”. The next sentence, 6:17, makes this clearer: “Who can go out of a house without using the door? Why does no one use this Way of mine?” Although the right Way seems so obvious to Confucius, it is not practiced at his time.<sup>12</sup>

---

9 Confucius, 2:8.

10 For further comments on 3:25, see Christian Helmut Wenzel, “Beauty in Kant and Confucius: A First Step”, *Journal of Chinese Philosophy* 33 (Jan 2006), 95–108, here 98–9.

11 For an interpretation of both 1:12 and 4:1, see Wenzel “Beauty in Kant and Confucius: A First Step”, 99–102.

12 *The Analects of Confucius: A Philosophical Reading*, trs. Roger T. Ames and Henry Rosemont, Jr. (New York: Random House, 1998), refers to the Dingzhou text and gives another translation of 6:16, preferring “humaneness” to “eloquence”. But I think we do not need to make such an adjustment. Confucius is not



How far can we go with our talk of the “inner” in Confucius? How far are we justified in doing so? And how far can we go with our talk of “morality” (*Moral*) and “beauty” (*Schönheit*), in Kant’s sense of these terms, in our interpretation of Confucius? If we say, as I just did, that a moral attitude—of filial piety, say—has to “shine through” and to give an act or a ritual performance its “beauty”, Kant would not agree.<sup>13</sup> Firstly, to him, within his transcendental philosophy, only objects of the outer senses can be called “beautiful”. Secondly, if a performance was beautiful, the grounds for this could not be moral ones, such as moral attitudes that somehow “shine through”. Beauty, Kant insists, must stand on its own feet. It must have its own specifically aesthetic justifying grounds, independently of morality. The “free play” of imagination and understanding must not rely on moral considerations, not even on *similarities* to moral reflections. The similarity must be gratuitous. Only then, Kant thinks, is it the case that beauty can serve morality. The link between the two is based on an essentially *gratuitous* isomorphism between reflections underlying judgments of taste and reflections underlying moral judgments. Aesthetic universality cannot be reduced to moral universality, in judgment as well as in reflection. Something cannot be beautiful because it is good, nor can it be good because it is beautiful.

But maybe we have already made a mistake at the beginning, by identifying *ren* with morality (*Moral*) and the inner, and by identifying *mei* with beauty (*Schönheit*).<sup>14</sup> Firstly, morality is a rather abstract notion in Kant. It belongs to reason: practical reason (*praktische Vernunft*). Kant lived at the time of the Enlightenment that stressed autonomy and rational abilities, and already in Aristotle we find strong links between morality and rationality, because theory and theoretic contemplation

---

happy with his time and does not approve of the situation where one unfortunately often needs eloquence as well. He does not recommend eloquence.

- 13 For a discussion of the relation between *ren* and *li* in general, not necessarily in relation to beauty, or *mei*, see Tu Wei-Ming, *Humanity and Self-Cultivation*, and Shun Kwong Loi, “Rén 仁 and Lǐ 禮 in the *Analects*”, in: Bryan W. Van Norden, *Confucius and the Analects: New Essays* (New York: Oxford University Press, 2002), 39–52. Tu emphasizes the creative, dynamic tension between the two. Shun discusses two extreme positions, one seeing *li* as defining *ren* and thereby making the latter depend on the former, the other taking *li* as merely instrumental for developing *ren* and thereby giving *ren* a more independent status. Shun himself suggests an intermediate position.
- 14 On *ren*, see Confucius 1:3 and 3:8; on *mei*, see 3:8, 1:12, and 6:16.

(θεωρία) always figure in the background. Both Aristotle and Kant were interested in the natural sciences. Nothing like this can be found in Confucius. Secondly, Confucius did not consider problems with the conceivability of freedom of will versus physical determinism, especially as we find the latter against the background of Newton's physics. Confucius therefore did not have to venture toward something like transcendental philosophy in order to cope with the problem of free will. Thirdly, morality assumes a rather intangible character in Kant, and one might wonder whether this is so in Confucius as well. According to Kant, one is never sure whether one has really performed a morally good act or whether some hidden selfish interest was involved. The moral person has to be excluded from the deterministic, physical world. It shrinks to a mere point. Confucius also expresses some kind of reservation or doubt regarding our knowledge of humaneness (*ren*)—he often says things such as: “I don't know if he is humane” and “I don't know how he can be called humane.”<sup>15</sup> But he is referring to the humaneness of others, not his own. Whether he has doubts regarding himself, in his own case and in first-person perspective, is another question.<sup>16</sup>

In spite of these fundamental differences and the problems of commensurability they create—and I think they indeed do create such problems—I nevertheless think it is fruitful to introduce a central notion from Kant's third *Critique* into the *Analects*: the principle of purposiveness (*Zweckmäßigkeit*). The subjective principle of purposiveness is the most important “moment” of the four moments of the judgment of taste, and it is supposed to help bridge the gap between nature and morality in Kant. The principle of purposiveness appears in three “modes”, as I would like to call it: It appears *subjectively* in beauty, *objectively* in our understanding of organic nature, and *transcendentally* in the natural sciences. In all three modes we find, to put it very generally, something fitting something else, as if by chance, and without us being able to account fully for why we find things the way we do. In aesthetics, we can read such experiences as signs. Kant speaks of “hints” (*Winke*), telling us that we fit into nature and that our hopes to realize our moral ideas in this world are not out of place. This should sound familiar, in

---

15 Confucius, 5:8 and 5:18.

16 In 7:33, Confucius remarks “The title of sage or humane man—how could I dare lay claim to such?” But this is less an expression of doubt, and more a sign of modesty.

some way, to anyone having read the *Analects*. What is more, the Kantian principle of purposiveness can be seen as giving support to Confucius' trust in rituals, as I will explain in the following.

When judging an object to be beautiful, we find it purposive for a free play between imagination and understanding, a play that is, firstly, pure (i. e., free from personal, individual interests), and secondly, harmonious, as would be required for cognition in general. The play of the faculties transcends the individual and opens horizons for cognition and discovery.<sup>17</sup> For Kant, even the pleasure it gives rise to is based on this feature of transcending the individual. This is part of what makes it an *a priori* pleasure. We contemplate and feel the pleasure as human beings in general, as human beings qua human beings. The symbolic link with morality is then based on this feature of generality; for Kant this is not mere empirical generality but even *a priori* universality. But there is more. On an even higher level, we take an "intellectual interest" in beauty by taking beauty as a hint that tells us that our moral hopes might not be frustrated.<sup>18</sup>

Talk of intellectual interest in beauty applies primarily to beauty of nature and not to beauty of art. But then, as the latter is the product of genius and genius is a gift of nature, both the hint and the intellectual interest can be found in beauty of art as well. This is relevant to our discussion of Confucius, because we can ask how "natural" a ritual possibly can be. We can mistrust rituals, pointing out that they are artificial creations of humans and that they can go wrong. This worry, in turn, could be countered by saying that ritual can be the product of genius, where genius is, in a Kantian way, inspired by "nature", or, in a Confucian way, by Heaven (*tien*). Nevertheless, we then still have the problem of determining in particular situations what is, and what is not, an act or suggestion of genius, or Heaven.

Confucius emphasizes ritual, because he wishes to re-establish order at a time of war and the disappearance of traditional values. Ritual has a practical purpose. But it is also harmonious (*he*) and beautiful (*mei*). This harmony is mainly social harmony, but I think it also has a cosmological

---

17 See Christian Helmut Wenzel, "Beauty, Genius, and Mathematics: Why Did Kant Change His Mind?", *History of Philosophy Quarterly* 18 (Oct 2001), and Wenzel, *An Introduction to Kant's Aesthetics*, 133–40, for an account of beauty and genius in mathematics within the Kantian framework. Even in mathematics, I think, some kind of freedom matters.

18 For Kant on our "intellectual interest" in beauty, see 5:298 (§42).

dimension. Although Confucius is mainly aiming at inter-subjective uniformity between human beings, as well as conformity with the traditional rules, time and natural circumstances have to be taken into account as well. In Kant, it appears, *prima facie*, very differently: the harmony in question here, aesthetic harmony, is a harmony between our cognitive faculties, imagination and understanding. This harmony is in us, in our ways of perceiving, understanding, conceptualizing, and forming ideas. In Kant it is an *intra-subjective*, in Confucius a more *inter-subjective* harmony that matters—at least so it might seem so far.

Of the three modes of purposiveness (subjective, objective, and transcendental), only the first is linked to beauty. The other two are related to teleology and the empirical natural sciences. Only the first is helpful in seeing beauty as a symbol of morality, whereas the other two modes point toward nature. Nevertheless, the element of chance and gratuity can be found in all three. We find ourselves fitting into nature on a gratuitous basis, and this again matters for Kant's concept of beauty.<sup>19</sup> How does Confucius fare in comparison with this?

Compared with *he* and *mei*, Confucian *tian* and *dao* figure more cosmologically. Although *dao* should be realized and practiced by us, it exists already in nature outside. Human harmony and beauty are derived from it. What comes closest to this in Kant might be the principle of purposiveness. Although this principle is *a priori* and part of our faculty of judging, and therefore in us (as is typical for transcendental philosophy), it nevertheless forms a bridge between nature as being given and nature as being systematically understood. It underlies empirical order (the second and third modes of purposiveness). Kant's principle of purposiveness thus offers a possibility for an explanation of the role of *tian* and *dao* in Confucius. It can be seen as giving grounds for trust in rituals. Of course, again, there is no guarantee. Rituals can go wrong.

The harmony Confucius is aiming at is a harmony of *dao*, and we can find this harmony in three ways: (1) between human beings, (2) in relation to traditional values and rules, and, though possibly to a lesser extent, (3) with nature as it is given. The first (morality) and the last (nature) can be seen to figure also in Kant's notion of purposiveness. But

---

19 I do not know how much Kant would have changed his views had he known of the evolutionary theories that we know today. He certainly foresaw many of their features, and it is remarkable that he interpreted the "beauty of nature" as being based on a "blind" and mechanical, and not a purposefully designed, nature (see the recent work of Alexander Rueger on this matter).

the second (tradition) we find only in Confucius and not in Kant. Confucius had faith in the past; Kant did not.<sup>20</sup> Kant trusts more in reason and the future, whereas Confucius believes in the ways of the past and their power to re-introduce *dao*. Confucius models his ideas on a concrete, though past, reality. Kant proposes abstract principles for the future.<sup>21</sup> Confucius trusts in ritual and the idea of the gentleman as a model for emulation. It is therefore not surprising that aesthetics and ethics are more closely related in his world view than in Kant's: aesthetics is about what we can see, or otherwise perceive, in the Greek sense of the word "aesthesia" (αισθησις), and in some sense we can see and perceive ritual performances and the behavior, manner, conduct, and comportment of a gentleman. We can see a model such as a gentleman or a ritual performance in time and space, but we cannot see the categorical imperative. The latter is abstract and involves reflection about maxims and their universalizability. Kant's interest in mathematical laws of nature had an effect on his conception of morality, placing him in a position more distant from that of Confucius. In his aesthetics

- 
- 20 For a brief presentation of views by Fingarette, Hall and Ames, and Roetz that do not see Confucius as really turning to the past, see Ivanhoe, *Ethics in the Confucian Tradition*, 5–10. This would count against the view presented here. Furthermore, Kant also saw values in the traditions from the past. He recommended setting up standards for poetry in Latin, because this language does not change any more and therefore provides stability. He also recommended discipline over genius whenever the latter tends to be exaggerated and superficial. But I think his trust in the progress of the sciences and in the power of reason and autonomy was stronger than his love for the past, and here I see the difference between him and Confucius, in whose views we do not find the natural sciences play such a role and have such a driving force.
- 21 No wonder Fingarette finds no "crossroads" in Confucius. There is only one past, whereas there are many possible futures. The past is already there and cannot be changed, while the future is open (at least so it seems, a determinist might say). Thus if one is oriented toward the past, there will be no "crossroads", while regarding the future one has to make choices and to set up maxims for oneself to act upon. It must be admitted, though, that when looking into the past for orientation, one often finds more than one instance that one could take as a model and choose to follow. Thus one still has to reflect and to make choices. For the "Way without crossroads", see Herbert Fingarette, *Confucius—The Secular as Sacred* (New York: Harper & Row Publishers, 1972), ch. 2. To avoid a wrong impression, it must be said that Fingarette sees Confucius as an innovator and not as turning to the past (ch. 4). He also thinks we should not look for an "inner" in Confucius (ch. 3). However, as much as I find his views interesting and thought provoking, I do not follow him in these two points. Compare note 1, above.

on the other hand, he took a position different from the rationalist tradition, inviting the idea that his views are closer to Confucius' in this domain. Thus aesthetics might appear to be a more suitable domain of comparison than ethics.<sup>22</sup> But then the element of freedom, introduced and insisted on in Kant's aesthetics, standing as it does in opposition to the German rationalists' attempts to introduce objective rules of taste, must be squared with Confucius' views of rituals. This is likely not to be so easy, because with Kant's emphasis on freedom in his aesthetics, it is easier for him to move to modern art for instance, than is the case with Confucius.

Of the three elements of harmony in Confucius, mentioned above, only (2), the element of beauty through harmony with a past reality and of following ritual to re-introduce *dao* in our human activities, is absent in Kant. Instead, we might say, we find in Kant (2') the *a priori* principle of purposiveness; this is not about the past or about *dao*, but about nature as given. This principle explains beauty and bridges what appears to be a gap between morality and nature. Such a gap is absent in Confucius from the start, because no natural sciences had torn nature and morality apart. No Newton and no Galileo had created the threat of physical determinism against our feeling of moral freedom and agency. The absence of this threat for Confucius, and the absence of the views this threat gave rise to, creates the biggest difference between the two thinkers' conceptions of morality. This is why (2) is very different from (2'). For Kant, there must be two separate worlds, one of moral freedom and another of physical nature, the noumenal and the phenomenal, to escape this threat. At least there must be two very different perspectives. Not so for Confucius. In his view, we find *dao* penetrating everything, outer nature as well as humanity in us. The inner therefore has a very different flavor in Confucius.

Kant and Confucius take it as evident that morality and *ren*, respectively, are valuable and that one cannot argue for their value. Nevertheless, there remain differences between them. Kant gives an explanation for beauty through his analysis of the judgment of taste, but Confucius does no such thing. Kant also tries to explain how beauty is linked to morality, whereas Confucius does not. Confucius simply "relies" on the link, as we might say when applying Kant's theory. He recommends

---

22 For a discussion of Kant's emphasis on the role of feeling in aesthetics, in opposition to the rational tradition of his time, see Wenzel, *An Introduction to Kant's Aesthetics*, 4–7.

the practice of rituals, not only pragmatically, counting on its order-creating function, but also aesthetically, relying on the harmony and beauty this practice creates. His intuitive idea that harmony and beauty “shine through” can be explained and justified if we have theories of aesthetic and moral reflection at our disposal as we find them in Kant. A further aspect of comparison between the two philosophers on the relationship between morality and aesthetics is that Confucius saw aesthetic qualities as refinements of moral ones, while Kant’s ideas can be applied, as in Schiller,<sup>23</sup> to develop the concept of an “aesthetic education”. But this is a topic for a separate essay.

The *a priori* principle of purposiveness even offers a guarantee of some sort, not a guarantee for each individual case, but a guarantee in general, regarding the possibility for such a link between beauty and morality. For Kant, this was important, because it creates grounds for our hope for a better future and it involves not only aesthetics and morality, but also a science-oriented metaphysics. It involves not only subjective and objective, but also transcendental purposiveness. The latter can be seen only very dimly in Confucius, if at all, because there is no theory in Confucius comparable to the Kantian transcendental theory of *a priori* purposiveness. Again, there simply was no Leibniz or Newton in China. Even Aristotle’s notion of *ἡεωρία* is far away from Confucius’ thoughts. On the other hand, Confucius may not have been a transcendental realist either. His views of *tian* and *dao*, expressed sometimes with hesitation and sometimes with affirmation, can be seen as expressions of idealist intuitions—I do not mean skeptical intuitions about the existence of the external world as we find them in Berkeley, but positive intuitions about human values as we find them in Kant. The *a priori* principle of purposiveness can therefore be seen as a principle allowing us to explain the role of *tian* and *dao* in such intuitions in Confucius.<sup>24</sup>

---

23 Friedrich Schiller, *On the Aesthetic Education of Man in a Series of Letters*, tr. Elisabeth M. Wilkinson and L. A. Willoughby (Oxford: Clarendon Press, 1982[1967]).

24 I wish to thank my students from National Chi Nan University, Taiwan, for their questions and discussions during a course I gave on aesthetics in Kant and Confucius, and I wish to thank Philip J. Ivanhoe, Cheng Chung-Ying, Mihaela C. Fistioc, Shun Kwong-Loi, and James Peterman for pleasant conversations and helpful comments on earlier versions on this essay.

## 21. China, Nature, and the Sublime in Kant

*Eric S. Nelson*

### 1. Introduction

Naturalistic and anti-humanist accounts of early Lǎo-Zhuāng (老莊) Daoism and of the uncanny or terrifying sublime suggest that the everyday life and conventional personality of the individual is interrupted and displaced by overwhelming impersonal powers that reveal the “human” to be a false construction and the world an aesthetic, natural, or mystical play of forces.<sup>1</sup> This is often portrayed as entailing an either/or between anthropocentric humanism, with all of its questionable assumptions about “the human” as distinct from animals and the natural world, and an impersonal naturalism that seems to depersonalize and de-individuate the person.

I will examine whether there is an alternative to both of these one-sided perspectives and argue that human beings can be individuated within and in the context of their natural world. Such a natural and yet still ethical individuation can be glimpsed in the work attributed to the ancient Chinese thinker Zhuāngzǐ (莊子), the *Zhuāngzǐ*,<sup>2</sup> and in Immanuel Kant’s *Critique of Judgment*. According to standard readings,

- 
- 1 I would like to thank Dan Lusthaus for his comments and challenging questions that have helped improve this essay. Note that I adopt the terms “Early” or “Lǎo-Zhuāng” Daoism for the sake of convenience. This convention does not imply that there is an underlying unity or school between them nor does it necessarily include or exclude other varieties of Daoism.
  - 2 The *Zhuāngzǐ* (莊子) passages are cited by chapter. I have consulted the *Library of Chinese Classics* Chinese-English edition of the *Zhuangzi* by Wang Rongpei, Qin Xuqing, and Sun Yongchang (Changsha: Hunan People’s Publishing House and Foreign Language Press, 1999); Burton Watson’s *The Complete Works of Chuang Tzu* (New York: Columbia University Press, 1968); Martin Palmer’s *The Book of Chuang Tzu* (London: Penguin/Arkana, 1996); A. C. Graham’s *Chuang-Tzu: The Inner Chapters* (Indianapolis: Hackett, 2001); Hyun Höchsmann and Yang Guorang, *Zhuangzi* (New York: Pearson Longman, 2007); and Brook Ziporyn, *Zhuangzi: The Essential Writings* (Indianapolis: Hackett, 2009).



this is a hopeless strategy to the extent that Kant is concerned with the person's transcendence and Zhuāngzǐ with its natural immanence. Further, these radically divergent texts have no shared language, kinship, or identity.

Instead of advocating a hidden affinity, a critical reading of both reveals that Kant's third *Critique* goes beyond his more typical complicity with the anthropocentric domination of nature and that the *Zhuāngzǐ* does not eliminate individuality and the human in its skeptical challenging of conventional human perspectives and concern with *dào* (道) and *tiāndì* (天地, heaven and earth, or "nature"). Between Kant and Zhuāngzǐ, there is an open or empty space for considering individuality in the context of the natural world.

This essay relies on a reinterpretation of the *Zhuāngzǐ* as an ethics of natural yet not inhuman individuation and responsiveness in order to examine Kant's problematic interpretation of China and its "mysticism", involving a troublesome racial-aesthetics, and Kant's articulation—more evocative of early Daoist approaches to nature and Chinese aesthetics—in the *Critique of Judgment* of nature as free natural beauty and the sublime.<sup>3</sup> By stressing human responsiveness to free natural beauty, Kant proves there is more than the human domination of nature as either: (1) a constituted product or (2) mere objects of use and exploitation. Still, in the core of the third *Critique*, it appears as if the sublime reveals nature to be more than the human world only in the end for it to be lesser than human dignity. Kant's sublime risks endangering the person while disclosing the possibility of reaffirming the dignity of the individual in relation to the natural world. If that dignity is not affirmed, the person is overwhelmed in the adventurous or the grotesque. It remains to be seen if the awe and terror of the sublime is the possibility not of a dignity and vocation outside of the world but of renewed individuation in relation to the forces of nature.

---

3 I argue for the ethical orientation of early Daoism in E. S. Nelson, "Responding with *dao*: Daoist Ethics and the Environment", *Philosophy East West* 59:3 (July 2009), 294–316; and "Questioning Dao: Skepticism, Mysticism, and Ethics in the *Zhuangzi*", *International Journal of the Asian Philosophical Association* 1:1 (2008), 5–19.

## 2. Kant and China: Aesthetics, Nature, and Race

Deploying an aesthetics of race, or racial aesthetics, Kant attributed qualities to the various races of the world in his *Observations on the Feeling of the Beautiful and Sublime* (1764). Kant identified Asian Indians and the Chinese with the moral-aesthetic category of the “grotesque”, remarking of the latter (2:252):

What ridiculous grotesqueries do the verbose and studied complements of the Chinese not contain; even their paintings are grotesque and represent marvelous and unnatural shapes, the likes of which are nowhere to be found in the world. They also have venerable grotesqueries, for the reason that they are of ancient usage, and no people in the world has more of them than this one.

Earlier in the *Observations*, Kant described the category of the grotesque used in this passage, undoubtedly revealing a lack of understanding of Chinese practices and painting, as a gradation of the sublime (2:214): “Unnatural things, in so far as the sublime is thereby intended, even if little or none of it is actually to be found, are grotesqueries.” The initial examples thereof are: duels, cloisters, graves of saints; castigation, vows, monkish virtues; Ovid’s *Metamorphoses*; and the empty subtleties of Scholastic philosophy (2:214–5). The grotesque is correlated by Kant with the “weaker understanding” of the fantast and crank (2:222)—that is, with what he considered to be enthusiastic dabbling in fantastic fiction and/or the mystical, such as Ovid (whom Kant continued to quote in his works) and Emmanuel Swedenborg, and with the ritualism, scholasticism, and superstition he associated with pre-modern Catholic Europe.

Kant does not appear to have ever changed his negative stance toward what he considered Chinese. Nor did he ever share Leibniz and Wolff’s affirmative reception of various aspects of Chinese philosophy, politics, and ethics, where ideas and practices from China are seen as examples that can instruct modern Europeans.<sup>4</sup> Kant’s apparent hostility toward the non-European world is not limited to China. This has been explained by reference to increasing European colonial activity and the escalating disrespect for other ways of life as inferior and to be subjugated, although Kant was critical of colonization and slavery,

---

4 On the relation of Leibniz and Wolff to China, see E. S. Nelson, “Leibniz and China: Religion, Hermeneutics, and Enlightenment”, *Religion in the Age of Enlightenment*, vol. 1 (2009), 277–300.

as well as by Kant's problematic development of the discourse of race in a scientific or pseudo-scientific language.<sup>5</sup>

In his 1764 work, Kant is extending to the other peoples of the world categories that are simultaneously anthropological, moral, and aesthetic, and are first used to classify other Europeans, including the "phlegmatic" Germans. The stereotypes of the Chinese being overly refined and cunning, superstitious and ritualistic, are elucidated through the category of the grotesque. As marvelous and stultified, as somehow inhuman and unnatural, Kant is suggesting, without clarifying in detail, that it is a misrelation to the sublime. Whereas the sublime ought to evoke awe or terror, yet always returns the observer to the dignity and moral vocation of the human, gradations of the sublime such as the adventurous and the grotesque leave the self, and accordingly human dignity and moral personhood, lost in the powers of nature and tradition.

Kant repeatedly returned to the loss of the person in nature that he perceives in the East. Unlike Leibniz and Wolff's positive reception of China, and akin to Malebranche's condemnation of the Chinese for being Spinozistic, Kant's lectures on religion from the mid-1780s associate Asian thought with the mystical experience of nature, assimilating it to Spinoza (28:1052):

To expect this [e.g., divine participation] in the present life is the business of mystics and theosophists. Thus arises the mystical self-annihilation of China, Tibet, and India, in which one deludes oneself that one is finally dissolved into the Godhead. Fundamentally one might just as well call Spinozism a great enthusiasm as a form of atheism.

Such an atheistic mysticism or enthusiastic naturalism is incoherent according to Kant, since it breaches the transcendental separation between immanence and transcendence, the sensible and its conditions and the supersensible whereof nothing cognitively meaningful can be stated. Kant's depiction in this passage targets not only Buddhism but also Daoism, given his interpretation of its identification with the monstrous and grotesque in "The End of All Things". In language that partly evokes

---

5 Compare Robert Bernasconi, "Will the Real Kant Please Stand Up: The Challenge of Enlightenment Racism to the Study of the History of Philosophy", *Radical Philosophy* 117 (2003), 13–22; "Who Invented the Concept of Race? Kant's Role in the Enlightenment Construction of Race", in R. Bernasconi (ed.), *Race* (Oxford: Blackwell, 2001), 11–36.

the *rú* (儒) or Confucian disapproval of Buddhism and Daoism that probably informed his sources, Kant claimed (8:335):

From this [improper dabbling in the transcendent] comes the monstrous system of Lao-kiun [i.e., Lǎozǐ (老子)] concerning the *highest good*, that it consists in *nothing*, i.e., in the consciousness of feeling oneself swallowed up in the abyss of the Godhead by flowing together with it, and hence by the annihilation of one's personality; in order to have a presentiment of this state Chinese philosophers, sitting in dark rooms with their eyes closed, exert themselves to think and sense their own nothingness. Hence the *pantheism* (of the Tibetans and other oriental peoples); and in consequence from its philosophical sublimation Spinozism is begotten ...

In line with Christian ontotheology, Kant interprets the nothing and nothingness as primarily negative and pantheism as its celebration rather than as the affirmation of things and life in their immanent significance. Friedrich Nietzsche turned these two elements, nothingness and the self-affirmation of life in its immanence, against each other in his critique of Buddhism and the Asiatic. Ironically, Kant's portrayal of the Chinese was applied to Kant in Nietzsche's abuse of him as "der große Chinese von Königsberg" and "das Königsberger Chinesenthum"; with such labels Nietzsche seems to have some combination of moralism, mysticism, and ritualism in mind.<sup>6</sup> Nietzsche's polemical identification is obviously insufficient either to excuse Kant or link Kant and Chinese philosophy in any serious way. Even if Kant had more knowledge of Chinese thought, he might have further identified Daoism with the fantastic, akin to Ovid, Spinoza, or Swedenborg, and Confucian philosophy with the ritualism and scholasticism of Catholicism. Such associations are not unfamiliar in some recent and better informed authors who should know better.<sup>7</sup>

Instead of concluding with Kant's questionable judgments about the Chinese, or the affinities Nietzsche intimates, the following sections concern the relation between the human and the natural by reexamining the significance and import of (1) the beauty of "free nature" and of the sublime in Kant's philosophy and (2) what evokes free natural beauty and the sublime in the *Zhuāngzǐ* and, to a lesser extent, the *Dàodéjīng*.

6 Friedrich Nietzsche, *Sämtliche Werke: Kritische Studienausgabe* in 15 Bänden (KSA), ed. Giorgio Colli and Mazzino Montinari (Berlin: de Gruyter, 1980): KSA, 5.144; KSA, 6.177; compare S. R. Palmquist's discussion in "How 'Chinese' was Kant?," *The Philosopher* 84:1 (Spring 1996), 3–9.

7 On the problematic assumptions of Western sinology, see the preface to Russell Kirkland, *Taoism: The Enduring Tradition* (London: Routledge, 2004), xi–xx.

In the next section I inquire into whether the third *Critique* can be interpreted as a middle ground between impersonal nature and moral personality. In the last section, I consider whether Lǎo-Zhuāng Daoism truly dissolves the human into mystical nature, as both Kant and Confucian critics of early Daoism contend. Rather than being mystical absorption in the static unity of the one, it might well indicate the possibility of independent ease within changing nature, just as Kant wrote of English gardens and Baroque design (5:126)—although cultivating natural worldly freedom cannot be said to be merely a project of the imagination in early Lǎo-Zhuāng Daoism—by “[pushing] the freedom of the imagination almost to the point of the grotesque, and [making] this abstraction from all constraint by rules the very case in which taste can demonstrate its greatest perfection in projects of the imagination.”

### 3. A Daoist Reading of Nature in Kant’s Third *Critique*

Kant’s anthropological speculations concerning the Chinese are part of the dubious development of Enlightenment discourses about race, yet his depictions of the grotesque and the sublime and of absorption into the inhuman persist as questions, given the continuing significance of Kant’s thought and contemporary debates concerning the actuality, import, and value of the idea of the human person. Kant’s impoverished assessment of Daoism remains to some extent recognizable in newer approaches that celebrate or fear the loss of the person. The works attributed to Lǎozǐ and Zhuāngzǐ continue to be associated with tendencies appearing to deny the moral personality of the individual. These tendencies include the mystical, the naturalistic, the anti-humanistic, and according to detractors whom I have responded to elsewhere, even inhumane and totalitarian government.<sup>8</sup>

The *Zhuāngzǐ* in particular is a work full of stories of fantastic transformations that undermine constant identity and threaten moral dignity and responsibility, celebrating the anarchistic and aesthetic playfulness of life and being free and at ease in the world, as well as philosophical dialogues and reflections that have become a focal point for discussions of skepticism and deconstruction that reveal conventional human action,

---

8 I describe and respond to a number of such criticisms by elucidating early Daoism’s ethical dimension in “Questioning Dao”, 5–19, and “Responding with *dao*,” 294–316.

knowledge, and values to be uncertain. In the context of post-humanist and postmodern interpretations of the uncanny and terrifying sublime, and of mystical and deep ecological approaches to nature, challenging the metaphysical assumptions that privilege the human in anthropocentric humanism and personalism, both skeptical and mystical depictions of early Daoism are interpreted as implying that the everyday personal life of the individual is interrupted, dismantled, and undermined or transformed by overwhelming and/or more elemental impersonal powers.<sup>9</sup> The person and the human are accordingly revealed to be artificial constructions, with the world being an aesthetic, natural, or mystical play of inhuman forces.

Despite the third *Critique* and the *Opus Postumum*, Kant's philosophy of nature was criticized in German Romanticism for neglecting the vitality and holism of nature and, in works such as Adorno and Horkheimer's *Dialectic of Enlightenment*, for deepening the instrumental domination and exploitation of nature.<sup>10</sup> For such authors, the *Critique of the Power of Judgment* is too little, too late, or simply inconsistent, in addressing nature as a vital interdependent whole or in an environmentally sensitive way.<sup>11</sup> Insofar as nature receives value in the end only in relation to human feeling, thought, and dignity, and thus has no independence in relation to the human, it is clear which side Kant falls on in the conflict between an anthropocentric humanism that values the person at the expense of the natural world and animal life, and an impersonal naturalism (whether scientific, romantic, or mystical) where the person disappears as transient part or fabricated composite.

---

9 Note that Jean-Francois Lyotard rejected the overly simplistic identification of the sublime with a politics of the sublime that would be terror and its celebration in *The Postmodern Explained: Correspondence, 1982–1985* (Minneapolis: University of Minnesota Press, 1993), 67–71.

10 Theodor Adorno and Max Horkheimer, *Dialectic of Enlightenment*, tr. Edmund Jephcott (Stanford: Stanford University Press, 2002), and this critique continues in Adorno's later works such as *Beethoven: The Philosophy of Music*, tr. E. Jephcott (Stanford: Stanford University Press, 1998), 80; *History and Freedom*, tr. R. Livingstone (Stanford: Stanford University Press, 2007), 209–10.

11 As, for instance, in Ann A. Pang-White, "Nature, Interthing Intersubjectivity, and the Environment: A Comparative Analysis of Kant and Daoism," *Dao: A Journal of Comparative Philosophy* 8 (2009), 61–78. As implied by the argument of this essay, I slightly disagree with the conclusion that nature in the third *Critique* is inconsistent or that Kantian transcendental subjectivity and worldly phenomenism are more compatible with early Daoism and environmental ethics.

The sensual-material or naturalistic moment is to some degree recognized in Kant's Critical philosophy: first, in the first *Critique's* principle of phenomenality that, however, leads back to the transcendental power and unity of consciousness; second, in the empirical and causal motivations that the person ought to overcome through the moral law; and, third, in the sublime that risks destroying the person while disclosing the possibility of reaffirming the dignity of the individual in relation to the natural world. By placing it at risk, the abyss and terror of the sublime heightens the feeling of life (*Lebensgefühl*) and, through its temporary interruption, the mind's own striving is shown to break with its absorption in sense-objects and surpass "every measure of the senses" (5:250).

Kant's account of the feeling of life is historically connected with early modern discourses of *vis viva* in Leibniz and the more materialist notion of the *conatus* in Hobbes and Spinoza. These concepts concern individuation in relation to the forces of nature, which for Kant is a question of moral sensibility and vocation. In relation to the forces and conditions of life, humans find their own purpose in themselves and individuate themselves as moral beings in a worldly context. Whereas the beautiful "carries with it directly a feeling of life's being furthered" (5:245), the sublime "is a pleasure that arises only indirectly; it is produced by the feeling of a momentary inhibition of the vital forces followed immediately by an outpouring of them that is all the stronger." Such moral individuation in response to nature is not the subsumption of a particular under a universal category or the exemplarity of a type, as with determinate judgment, and thus not the pure dominion of active spirit over passive nature.

Instead of being the assimilative drive and mastery of the self-interested *conatus*, as some critics have interpreted *Lebensgefühl*, it is the undetermined responsive and reflective generation, formation, and cultivation of individual and social aesthetic and moral sensibilities in relation to particular phenomena. The feeling of life is the possibility of a prereflective awareness of self and other. Without a predetermined concept, it involves the nexus of nature as significant in itself and human feeling that cultivates nature's significance, even if sensibility must transcend the senses and sensuality to realize its rational vocation for Kant. The third *Critique* is not only a work about the generation and articulation of concepts. It concerns the coming to word and concept of what is heterogeneous, not given, or without a concept (20:202–3):

the sensuous, the natural, and the felt in art and genius, language, and the *sensus communis*.

The *sensus communis* is a sharing of sense proceeding through feelings rather than a common understanding working through concepts (5:238, 293). It is without a determinate concept or judgment and is universally communicable in requesting assent as distinct from legislating agreement (5:216, 221). As such, it allows for the interpretation and communication of the non-cognitive and non-conceptual, particularly feeling.<sup>12</sup> It constitutes a field of preunderstandings that do not only have the conservative function of reproducing custom, habit, and tradition, since it is communicative and can be open to and transformed by the new. The height of individuation in Kant is the genius who discovers ideas and ways of expressing them (5:317). Although restricted by the demands of rationality unfolded in the three *Critiques*, genius provides new forms and models for encountering and interpreting phenomena and oneself, as the genuinely and transformational “otherwise” has an important role in approaching society, culture, and art.

In the *Critique of Judgment*, Kant explored how nature can be judged reflectively as having purposes, humans can be said to be ultimate purposes, and art can embody and enact “purposefulness without purpose” as “lawfulness without law” or purposiveness without a concept of a determinate teleological final cause. This playful and anarchic removal of barriers and predetermined purposes in experiencing the free-form of the beautiful—for example, what is free in not being grounded in the concept of how the object ought to be or in an idea of purpose or perfection (5:222, 229)—and the formlessness of the sublime is connected with the feeling of life and contrasted with the seriousness of ethical, political, and religious purposes as governed by fixed forms and final ends.

The “unison in the play of the powers of the mind” is not a confused concept or inadequate idea but a feeling of inner sense (5:228). Such felt spontaneity and playfulness, as the promise of freedom from a predetermined purpose and as responsiveness in relation to the forces and conditions of life, indicates a non-instrumental, non-coercive, and non-dominating activity understood as a creative receptiveness or responsive spontaneity in encountering the myriad things and the world as an ineffable whole inviting further investigation and inquiry.<sup>13</sup>

---

12 Rudolf Makkreel, *Imagination and Interpretation in Kant* (Chicago: University of Chicago Press, 1990), 164.

13 On the spontaneity and responsiveness of life, note Makkreel, 106, 156.



Kant's third *Critique* does not, therefore, purely defend either the aesthetic use or moralistic domination of nature, as critics such as Günter Wohlfart contend. It is deeply ethical in locating the individuation and self-articulation of the person in a worldly, sensuous, and bodily as well as a social context.<sup>14</sup> The self does not cognitively or theoretically know itself. Instead the self finds itself, according to the third *Critique*, in its comportment, cultivation (*Bildung*), and culture in relation to nature, the sublime, and the supersensible.<sup>15</sup> As argued by John Zammito and others, this work was partly Kant's reply to the pantheism controversy that impacted German intellectual life in the late 1780s.<sup>16</sup> As opposed to being primarily reactive against pantheism and early romanticism, both clearly rejected by Kant, the third *Critique* articulates an alternative or middle ground affirming the person in relation to the forces and conditions of nature through feelings of life such as those of the beautiful and the sublime. Whatever their relation to the human faculties, Kant insists that both the beautiful and sublime please intrinsically, for themselves, rather than instrumentally, for something else, and that their purposiveness cannot be reduced to purposes—that is, instrumentally to human purposes.

Kant's approach to the feeling of life, reflective judgment, and *sensus communis* in the *Critique of Judgment* are ways of non-mechanistically, yet not metaphysically or teleologically in the strong sense, experiencing and articulating the nexus of life. They are non-mechanistic insofar as Kant describes the beautiful as free of calculative and instrumental interest, and the sublime as contrapurposive, addressing nature through a reflectively articulated purposiveness without a predetermined purpose. This nexus of life involves both the "external" natural world and the "internal" relations of the faculties of the subject.

---

14 In writings such as *Die Kunst des Lebens und andere Künste: Skurrile Skizzen zu einem eurodaoistischen Ethos ohne Moral* (Berlin: Parerga Verlag, 2005), Günter Wohlfart contends that Kant's philosophy is tied to an individualistic domination of nature and Daoism liberates us from such problematic individualistic humanism. For a condensed version of Wohlfart's argument, see his essay in the present collection (Keynote Essay 2). If my argument is correct, Kant's thought is more open to nature and Daoism to ethics and the individual than Wohlfart contends.

15 Kant, 5:265; on the sublime and supersensible conditions of the subject and its moral cultivation, compare Makkreel, 79–81, 83–4.

16 John Zammito, *The Genesis of Kant's Critique of Judgment* (Chicago: University of Chicago Press, 1992), 6–12, 228–48.

The analysis offered here places Kant's Critical philosophy in a different light, as having a "hermeneutical" dimension insofar as the human subject intrinsically lacks the transparency of self-knowledge (at least as a rationalistically intelligible essence) yet does live from the feeling of life that opens up questions of the self-understanding, interpretation, and individuation of that life. The proto- or quasi-hermeneutical aspects of the third *Critique* do not suspend or escape the conditions and demands of theoretical and practical reason. The book indicates strategies for a hermeneutics of "a" life or "individuated" life that do not rely on metaphysical self-knowledge or the rational psychology of the soul.

Kant's thought does not leave us with the bare mechanistic nature of the natural sciences nor return to a metaphysical or strong teleological conception of nature. It addresses questions of the formation and individuation of personal identity through reflective judgment and the *sensus communis*; these do not command or legislate to the phenomena but unresponsively or responsively interpret and communicate with them, in the context of the heightening and lessening of the "feeling of life" that seeks a balance and harmony in relation to itself and its world.<sup>17</sup> Such dynamic harmony does not deaden the mind with a static unity, since it is animated and enlivened with the connections and resonance between what is different and singular (cf. 5:219).

The singular "this" indicated and addressed in feelings and judgments of taste—"this rose is beautiful", to use Kant's example—is distinct from the general or universal spoken of in logical judgments, including those that are aesthetically oriented, such as "roses are beautiful" (5:215). Whereas one evokes the experience of dynamic harmony with a particular, without subsuming it under a pre-given concept insofar as the concept is in need of being articulated, the other subsumes or synthesizes particulars according to a predetermined concept. As Makkreel notes of Kant's distinction, such harmony is a balancing instead of a determinate synthesis or totalization: "A harmony involves a reciprocal relation between two distinct elements; a synthesis, as Kant conceives it, involves a one-sided influence for the sake of a strict unity."<sup>18</sup>

Kant's thinking of harmony in a play of forces and conditions, including in the face of the terror of the sublime (with the human disposition rising above sense objects and beginning to realize its non-sensu-

---

17 Makkreel, 3–6.

18 Makkreel, 47.

ous and moral vocation), offers an alternative to: (1) an overly anthropocentric reading of Kant; and (2) Kant's own inadequate appreciation of Chinese painting, aesthetics, and early Daoism. In their own sensibility and language, the latter are about the harmony and balance within the individual (as the free, unforced balancing in play of flood-like *qi* (氣), in contrast to Kant's language of the faculties) and in the individual's relations with the enviroing natural world that it transcends, without abandoning, in responding to things with freedom and ease.<sup>19</sup>

#### 4. Nature and Freedom in Kant and Early Daoism

Kant associated the Chinese, Tibetans, and Indians with Spinozism. This indicates a lack of knowledge of both Spinoza and South and East Asian peoples and cultures—if not worse, since Bernasconi and others have concluded that Kant was systematically ethnocentric and racist.<sup>20</sup> Despite Kant's actual racism, however, his thought surpasses these problematic motives. Thus, for example, his moral thought is radically egalitarian in its scope, and—as shown in this essay—his thinking of the natural and the human is not as foreign to non-western ways of thinking as he himself thought or as critics such as Günter Wohlfart continue to argue.

In the *Critique of Judgment*, Kant explored both art and nature as embodying “purposefulness without purpose” or the form of purposiveness without a cognitively represented end or teleological final cause. This playful and an-archic lack of purpose is contrasted with the seriousness of ethical, political, and religious purpose, including the racial aesthetics and anthropology found in other works. Such spontaneity and playfulness, as freedom from a preordained purpose, can be analyzed in relation

---

19 Chapter four of the *Zhuāngzǐ* describes how it is by knowing without knowledge and by emptying the self through the “fasting of the mind” that one opens oneself to the spontaneous responsiveness of one's vital energy or force (*qi*), receiving in sincerity and generously responding without assertion or imposition. See Höchsmann / Guorang, 103–4, Palmer and Breuille, 29–30; Watson, 57–8; Ziporyn, 26–7.

20 In particular, R. Bernasconi, “Will the Real Kant Please Stand Up”, 13–22, and “Who Invented the Concept of Race”, 11–36.

to the image and model of “free and easy wandering” unfolded in the *Zhuāngzǐ*.<sup>21</sup>

Such an analysis does not reveal any determinate parallels and analogies. It does, however, allow for the reconsideration of whether Kant privileged the human and neglected the natural and whether early Daoism neglected the human in prioritizing nature. These radically divergent perspectives offer two different articulations of a non-instrumental, non-coercive, and non-dominating activity understood as either (1) wandering free and at ease in the world or (2) a creative receptiveness or responsive spontaneity in encountering the myriad things and the world.

Whereas Kant emphasized the non-conceptual yet universal satisfaction enacted in the non-attached and free play of forces in aesthetic judgment, the *Zhuāngzǐ* articulates a non-conceptual and non-attached play that involves transitions between a multiplicity of perspectives. This includes the contra- or counter-purposive that Kant finds displeasing in the beautiful, despite the role he gives it in the sublime and his noting the beauty of the useless and hence free object (5:210–1, 245). Instead of limiting this multiplicity and variability of transitions and perspectives to the freedom of the imagination and play in the aesthetic domain, and ultimately subordinating it to morality as in Kant’s third *Critique*, the Daoist sage is portrayed as responsively free and at ease amidst the myriad things.<sup>22</sup> *Zhuāngzǐ*’s responding without retaining, acting upon without harming, is more expansive than any conditional and limited goal or purpose that would limit the self to its perspective without recognizing its inherent transience and multiplicity. Such responsiveness does not—to speak Kant’s language—presuppose and is not restrained by a determinate concept, even though it employs concepts and words that are unfixed yet not therefore meaningless. Liberation from the determinate, purposive, and useful enables human beings to relate to things and their context in a fundamentally different, non-instrumental way. This way

21 Due to space limitations, I have been able to develop the Daoist part of my argument only schematically. I refer those who are interested in Daoism to my articles: “Questioning Dao”, 5–19, and “Responding with *dao*,” 294–316.

22 Hyun Höchsmann and Christian Wenzel have addressed such concerns by stressing the practical-ethical character of freedom in both Zhuangzi and Kant in their respective articles: “The Starry Heavens above—Freedom in Zhuangzi and Kant”, *Journal of Chinese Philosophy* 31:2 (June 2004), 235–52; and “Ethics and Zhuangzi: Awareness, Freedom, and Autonomy”, *Journal of Chinese Philosophy* 30:1 (March 2003), 115–26.

cultivates the self but is not therefore egotistical, since it calls on the self to individuate itself amidst things. It is naturalistic yet not thereby inhuman, if it is human to be oriented by and respond to heaven and earth.

Each of these works in its own way concerns individuation through cultivating balance in relation to nature within and outside oneself. Kant's third *Critique* and the *Zhuāngzǐ* are not simply aesthetic. They are deeply ethical works in (1) challenging the instrumental reduction of nature and the naturalistic reduction of the person or individual and (2) indicating the freedom in interaction and harmony between the human and the natural world. Nevertheless, despite such resonance, Kantian and Daoist visions of freedom and balance in relation to self and world remain incommensurable.

As Kant noted from his *Lectures on Ethics* to the *Critique of the Power of Judgment*, nature and animals are not to be purely instrumentalized, exploited, or treated with indifference or cruelty, since how humans relate to them reflects how they sense and cultivate their own feeling of life and moral vocation. In the context of the moral heightening and formation of the feeling of life, humans have indirect and mediate duties to animals, including negative duties against cruelty and positive duties of love and humaneness (27:459, 710), although such duties and sentiments are ultimately subordinate to the necessity of human needs (27:460). While Kant did recognize nature's beauty and sublimity independently of calculative interests and limited human—in contrast with moral-vocational—purposes, he still demanded the person's separation from nature for the sake of morality and the postulates of morality (freedom, immortality, and God), thereby rehabilitating theistic and transcendent religion, as argued in the third *Critique's* concluding pages. In contrast, although not without recognition of the transcendence or transformation within immanence, the Zhuangzian Daoist finds ethical independence dwelling within nature itself and disinterestedly embracing the myriad things in the immanence of their singular self-so-ness (*zìrán*): following each being's own grain, including one's own, and accordingly discovering one's freedom in the midst of the world.<sup>23</sup>

---

23 This difference evokes Kant's critique of the self-sufficiency of morality (that it is its own reward regardless of hope in a future life) and nature (that it exists indifferently in and out of itself without regard for human hopes) in Spinoza, with which Kant associated Daoism (5:452).

BOOK TWO:

CULTIVATING PERSONHOOD  
IN POLITICS, ETHICS, AND RELIGION

PART V

The Status of Persons  
in Applied Ethics and Law



## 22. Is There a Kantian Perspective on Human Embryonic Stem Cells?

*Ronald M. Green*

As my title indicates, my research on this topic began with a question. What does Kant's ethical theory tell us about the moral acceptability of human embryonic stem cell (hESC research)? This biomedical technology promises to deliver powerful new medical therapies, but it does so by requiring the destruction of early human embryos.<sup>1</sup> As a result, determining the moral status of embryos is crucial to the ethical assessment of hESC research. My title is also meant to suggest that my question remains unanswered. After reviewing the relevant loci in Kant's writings and others' interpretations of their meaning, I conclude that Kant's ethics is equivocal with regard to the moral status of human embryos, such that his ethics can be interpreted as either supporting or condemning human embryonic stem cell research. This conclusion may not be as helpful as the answer I searched for, but it illustrates the presence of an unresolved theoretical problem at the heart of Kant's ethics.

Kant's writings admit of at least two different ways of arriving at a determination of the moral status of early human embryos. Each answer points in different directions, and each leaves room for disagreement about how Kant might regard the ethics of hESC research.

The first approach leads to a conclusion that minimizes the moral claims made on us by embryos. It begins with the observation that while the second formulation of the categorical imperative requires us never to treat humanity simply as a means but always as an end, the basis for this dignity resides in our active ability to exercise autonomous reason. As Bertha Alvarez Manninen observes, "The term 'humanity' is not meant to denote a membership in any group at all, rather it denotes a certain capacity possessed by all persons, i. e., accountable beings, to

---

1 In fact, new technologies such as single cell blastomere biopsy and induced pluripotency may obviate the need to destroy embryos in order to produce a pluripotent stem cell line, but much research in this area continues to rely on cell lines derived from human embryos.



reason and set ends for themselves.”<sup>2</sup> Allen Wood terms this a “logocentric view” of human personhood because “it is based on the idea that rational nature, and it alone, has absolute and unconditional value.”<sup>3</sup> It follows that one need not be human, in the biological sense, a member of the species *Homo sapiens*, to be a person. Kant in fact speculates on the possible existence non-human rational beings on other planets (7:331). But if some moral persons are not human, neither are all entities that are biologically human moral persons. Since embryos do not actively exercise reason, they are not “persons”, undoubted possessors of dignity, the subjects and objects of moral respect.<sup>4</sup>

If we interpret Kant as attributing moral personhood only to beings capable of actively exercising reason, however, this does not necessarily lead to the conclusion that embryos lack all moral status and can simply be used as means to our ends, as hESC research requires. Kant maintains that we have duties even to non-rational entities, and to nature itself. In the *Metaphysics of Morals* and *Lectures on Ethics* he states that although we have no “immediate duties” to animals, we do have duties toward them that are “indirect duties to humanity” (6:443; 27:459–60). The abuse of animals, Kant believed, erodes human beings’ capacity for sympathy and compassion, with possibly dangerous implications for our treatment of one another. Although Kant believed we can use animals for essential human needs (such as labor or food), he insisted they should be slaughtered as quickly and painlessly as possible (6:443). Kant would also allow painful experiments on living animals if the ends were sufficiently important, but he strenuously objected to such experiments “for the sake of mere speculation”, or if their goal could be achieved in other ways (6:443).

It is not entirely clear what the implications are of applying the concept of “indirect duties” to human embryos. Some opponents of embryo research have sought to base their argument not on any basic moral rights of the embryo but on the possible impact of its mistreatment on our own moral character. If we wantonly instrumentalize and destroy this incipient form of human life, they maintain, we risk en-

---

2 Bertha Alvarez Manninen, “Are Human Embryos Kantian Persons?: Kantian Considerations in Favor of Embryonic Stem Cell Research”, *Philosophy, Ethics, and Humanities in Medicine* 3:4 (2008), 1–16.

3 Allen W. Wood and Onora O’Neill, “Kant on Duties Regarding Nonrational Nature”, *Proceedings of the Aristotelian Society* 72:1, Suppl. (1998), 189–228.

4 For a statement of this Kantian view as applied to embryos, see Louis M. Gueinin, “Morals and Primordials”, *Science* 292 (2001), 1659–60.

couraging the abuse and exploitation of whole classes of human beings: the infirm, the handicapped, or the elderly.

In reality, however, such claims are hard to sustain. Because the early embryo is composed of a small number of undifferentiated cells, it lacks sentience, the feature that underlies and explains our concern with animal suffering. Early embryos do not even possess the bodily form of the fetus, a fact anti-abortion activists have exploited by presenting gruesome images of aborted fetuses to marshal support for their cause. Finally, there is little evidence that most human beings identify emotionally with early embryos. Hundreds of thousands of frozen embryos are stored in infertility clinics in the United States.<sup>5</sup> Almost all of these will eventually be destroyed. The process of in vitro fertilization typically requires the creation of more embryos than can be safely transferred to a woman's womb and the subsequent discard of those that are left over. Yet, despite this vast wastage of embryos, most people appreciate IVF and are prepared to use it to overcome infertility, including those whose religious tradition opposes hESC research. Against this background, it is hard to see how the destruction of early embryos for hESC research is likely to degrade our current levels of respect for human life. I cannot imagine that Kant, who is so moving and so contemporary in terms of his concern for animal suffering, would place the protection of embryos over the real health needs of children and adults that could be met by hESC research.

Mention of children brings me to the second way Kant could arrive at a determination of the moral status of early human embryos. In this approach embryos are evaluated not in terms of the indirect effects of their treatment on moral persons, but as possible persons in their own right. The textual basis for this is Kant's treatment of the relation of parents and children in §28 of the *Rechtslehre*. Children, of course, are not autonomous rational persons. For most of their early development they lack a mature capacity for rational decision making and moral responsibility. Most legal systems recognize this in sparing them the punishments that adults receive for committing criminal acts. Yet Kant does not for this reason deny children the measure of moral respect due adults. "Children, as persons," he says, "have by their procreation an original

---

5 Rick Weiss, "400,000 Human Embryos Frozen in U. S. Number at Fertility Clinics Is Far Greater Than Previous Estimates, Survey Finds," *Washington Post*, Thursday, May 8, 2003, A10.

innate right to the care of their parents until they are able to look after themselves.” This is true, Kant continues (6:280–1),

For the offspring is a person ... So from a *practical* point of view it is quite a correct and even necessary Idea to regard the act of procreation as one by which we have brought a person into the world without his consent and on our own initiative, for which deeds the parents incur an obligation to make the child content with his condition so far as they can. They cannot destroy their child as if he were something they had *made* (since a being endowed with freedom cannot be a product of this kind) or as if he were their property, nor can they even just abandon him to chance, since they have brought not merely a worldly being but a citizen of the world into a condition which cannot now be indifferent to them even just according to the concepts of Right.

Kant’s clear attribution of personhood to children poses a serious problem for the view that Kantian personhood depends on the active exercise of rational autonomy. This has led some students of Kant’s theory of personhood to propose a different understanding of his conception of rational capacity, an understanding that also supports the widely held moral belief that children merit at least as much moral respect as adults. Thus, Allen Wood argues that in speaking of rational capacity in a Kantian framework we must make a distinction between those who are persons “in the strict sense” and those who are persons “in the extended sense.” The former category encompasses individuals who are presently capable of thinking and acting rationally, while the latter those who, in Wood’s words, exhibit “fragments” of rationality or “necessary conditions of it.”<sup>6</sup> This includes “small children and people who have severe mental impairments or diseases which deprive them, either temporarily or permanently of the capacity to set ends according to reason.”<sup>7</sup>

Wood’s account offers a compelling and textually well-grounded interpretation of Kant’s position, but, as Wood admits, it poses a question with important implications for the stem cell debate. “Exactly at what stages of human life should beings be regarded as persons in the extended sense?”<sup>8</sup> How far back into the human reproductive process does this “extended sense” of personhood reach? Does it go back to the fetus in utero? Answering “yes” to this has powerful implications for the question of abortion, for if parents have “no right to destroy [their child] as if

---

6 Wood and O’Neill, “Kant on Duties Regarding Nonrational Nature”, 198.

7 Wood and O’Neill, “Kant on Duties Regarding Nonrational Nature”, 198.

8 Allen Wood, *Kantian Ethics* (Cambridge: Cambridge University Press, 2008), 97.

it were their own property”, it is hard to see how a woman could have any more right to destroy the baby she is carrying. Does it reach back to the embryo in vitro? If so, it would be difficult to justify destroying embryos to produce hESC lines.

An oblique remark by Kant in the *Metaphysics of Morals* only sharpens these questions. In the course of discussing suicide, Kant says (6:422), “Willfully killing oneself can be called **murdering oneself** (*homicidium dolosum*) only if it can be proved that it is in general a crime committed either against one’s own person or also, through one’s killing oneself, against another (as when a pregnant woman takes her life).” Kant is not here discussing abortion, but the suggestion that killing a fetus might be regarded as homicide, even when the mother is the agent, suggests that he sees personhood as reaching back in Wood’s extended sense to the period before birth.

Wood himself does not draw this conclusion. “It is one thing to say that parents should be thought of as bringing a person into being, and even that they have duties of care to their offspring from conception”, he says. “It is quite a different thing to say that the offspring is a person from conception onward. The first two things Kant does appear to say; the third is something he never quite says.”<sup>9</sup> In a specific effort to rebut the claim that the embryo or fetus is a person, Wood points to the problem posed by the fact that a fetus resides in a woman’s body such that protecting it may run counter to her vital interests:

Regarding the question of whether an embryo in vitro is a person in the extended sense, that should turn on whether, in order to treat it as a person, some woman would have to be coerced into having the embryo implanted in her uterus and then compelled to carry it to term. Clearly if she would, then the embryo should not be judged a person in the extended sense. I conclude that if granting to embryos or fetuses the same “right to life” that is thought to belong to persons in the extended sense would involve such coercive or invasive conduct, then it would constitute gross disrespect to rational nature to grant them that status.<sup>10</sup>

A similar argument has been advanced by Susan Feldman, who sees privileging the fetus over the mother as a consummate illustration of “treating a person as a mere means.”<sup>11</sup> Manninen applies this argument

---

9 Wood, *Kantian Ethics*, 39.

10 Wood, *Kantian Ethics*, 98.

11 Susan Feldman, “From Occupied Bodies to Pregnant Persons: How Kantian Ethics should Treat Pregnancy and Abortion”, in Jane E. Kneller and Sidney

to the ex utero embryos used to produce hESC lines. Pointing to the nearly half million frozen embryos left over in fertility clinics across the United States, she observes that at least until ectogenesis becomes a reality and embryos can be gestated without the use of a woman's womb "treating these embryos as persons may entail implanting those embryos into the uteruses of women, whether or not they desire to gestate the embryos .... Such an action would disrespect the actual capacity for humanity that women possess."<sup>12</sup>

While I am sympathetic to the much higher valuation all these writers accord to women relative to embryos and fetuses, I do not think this argument warrants the conclusion that Kant did not regard prenatal entities fully as persons. If children are persons because of their incipient rationality—rationality in the extended sense—embryos and fetuses might be persons as well. If they are, why should the fact that their survival is in conflict with the vital needs of another person efface their personhood or reduce their moral claims on us? After all, vital conflicts often arise between human adults or between adults and children, and this does not readily justify active killing as a means of eliminating the conflict. No one would advocate resolving a situation of famine by killing some of the contenders for scarce food. Furthermore, even if it is true that some conflict situations have led to extreme measures (as in some of the classic lifeboat cases), these are typically desperate circumstances. At most, this would warrant abortion or embryo discard only when this is needed to prevent the imminent death of another person (the mother or stem cell recipient). This hardly justifies most abortions or the destruction of embryos for purely research purposes.

Other strategies have been adopted to try to work around the obstacle posed by these Kantian statements about children and fetuses. For example, Georg Geismann denies that the embryo or fetus's potentiality for rational development qualifies it as possessing moral personality because personality, for Kant, is not a matter of natural facts. Like the freedom it presumes, personality resides outside the realm of phenomena and cannot be founded on any features of a developing biological organism.<sup>13</sup> But Geismann overstates the matter. It is true that Kant offers a

---

Axinn (eds.), *Autonomy and Community* (Albany: State University of New York Press, 1998), ch. 13, 265–82.

12 Manninen, 10.

13 Georg Geismann, "Kant und ein vermeintes Recht des Embryos", *Kant-Studien* 95 (2004). 443–69. For a similar view, see Mark Sagoff, "Extracorporeal Em-

dualistic conception of human agency, seeing us as free despite the necessary causality of our biological nature, but he nowhere says our freedom and personality do not arise out of or somehow depend on our biology. Quite the contrary. In the same section of the *Rechtslehre* where he discusses the personhood of children, Kant observes the deep puzzle of human procreation: that free beings come into being as the result of a physical act. This reality is difficult to understand, but it need not be denied. In typical fashion, Kant asserts that while reason is not able “to make this relation of cause to effect comprehensible for theoretical purposes” it need not be dogmatically rejected (6:281n). “All that one can require of reason here would be merely to prove that there is no contradiction in the concept of a *creation of free beings*.”

If, therefore, freedom can and indeed must have a material foundation for creatures like ourselves, we are left with the question of what that foundation consists in and how much of it is required. Kant clearly believes that the being of young children warrants according them the status of persons. Is this true of fetuses? Perhaps. What about embryos? Opponents of embryo research using Wood’s “extended sense” of personality might point back to the human genome that comes into being sometimes around or shortly after fertilization as the appropriate material basis for personality. Unfortunately, there is nothing in the Kantian corpus that tells us whether they are right or wrong. Indeed, there is nothing in this corpus that even tells us what are the properties of children that make them persons.

To here I have presented two approaches to Kant’s treatment of personhood, one resting on indirect duties to ourselves, the other on a concept of “extended humanity”, that potentially lead to different estimates of the moral status of early human embryos, and with that, the permissibility of using embryos as material for the production of valuable hESC lines. The sustainability of each of these possibilities on the basis of basic Kantian insights and Kant’s own texts, tells me that we cannot arrive at a single, indisputable Kantian view on hESC research. This said, I want to conclude by offering my own depiction of where I believe Kant would come out if he were to examine this issue with the full information of modern biology and in a manner consistent with some of his deepest moral insights.

---

bryos and Three Conceptions of the Human”, *American Journal of Bioethics* 5.6 (2005), 52–4.

I begin with the undeniable fact that Kant views children fully as moral persons. We have seen that Wood attributes this to their rationality in the “extended sense” that includes their possession of “fragments” of rationality or “necessary conditions of it.” But I think Wood may be mistaken in placing his emphasis on properties possessed by a child. Rather, what seems to be most salient in Kant’s mind is the obligation that is owed to the child because it is both in our midst and on its way to being an equal member of our moral community. Parents have no right to destroy their child “nor can they even just abandon him to chance, since they have brought not merely a worldly being but a citizen of the world into a condition which cannot now be indifferent to them even just according to the concepts of Right” (6:281).

The focus here, it seems to me, is on an independently existing being whose abuse or neglect has obvious effects on us all. Not only does parental neglect force all of us to witness the mistreatment of beings very much like ourselves, but it also threatens to lead to the creation of adults who have been victims of neglect or abuse. This is hardly a recipe for preparing future members of a hoped for Kingdom of Ends. In other words, we must treat children as persons not so much because they possess some qualities related to rationality or the conditions for it (as Wood maintains), but because they normally become persons whose moral character we must protect and promote. The argument that we could eliminate this problem by permitting only the *killing* of children—the practice of infanticide but not abuse—is not convincing. Any cultivation of behaviors and attitudes of violence toward children threatens the safety and well being of all children.

This suggests to me that there are good reasons for treating children *as though* they are persons from the moment of their appearance among us as independent human beings. The concern to protect such beings is more than an indirect duty to ourselves: it is a very direct duty to beings that will become adults. I think it is reasonable and consistent with the relevant passage in the *Rechtslehre* to read Kant as saying just this. If I am right, then it is also unreasonable to believe that Kant, if he possessed the information provided by modern embryology, would regard the early embryo as a future citizen of the world whose life and character we also must protect. The destruction of embryos (in IVF or hESC research) does not tug at our heartstrings, nor will it produce psychologically or physically damaged adults. In addition, the great majority of early embryos naturally perish during the first days and weeks of development. This makes it hard to see why we should invest emotional con-

cern in a class of entities with which nature is so profligate. Once we shift our attention away from uncertain questions about which properties must be possessed by incipient human beings to qualify them as persons in an extended sense, and focus instead on the implications of mistreating a class of human beings, it becomes clear, from a Kantian perspective, that embryos are not likely to merit much concern.

Although I favor this interpretation of Kant's position, I confess that it is not the only one that is possible. Wood's conception of rationality in the extended sense is a reasonable interpretation of Kant's comments, as are those views that draw on the notion of rationality in an extended sense to give even more protection to incipient human life at the earliest stages of development than he does. Then, too, there is the curious passage in the *Metaphysics of Morals* that alludes to a possible double homicide when a pregnant woman commits suicide. If this represents Kant's view of the moral status of the fetus, it suggests a different approach to personality than the one I have presented for born children, since abortion has none of the future impacts on persons that child abuse or neglect do. The passage suggests that potential rationality alone justifies an ascription of personality. It reflects a view of rationality in the extended sense that is so strong that it might well reach back to the embryo. However, Kant's remark is cryptic. It is less an assertion than a hypothetical casuistic question. It also may be little more than a hyperbolic expression of Kant's well-established opposition to suicide. In any case, it would surely be perilous to found an entire Kantian theory of personhood and abortion on this remark.

On this note I will conclude. Although I have offered my preferred reading of Kant on the matters of the status of human embryos as it might bear on stem cell research, I have shown that the texts remain equivocal. Despite the absence of closure here, I think this finding is positive. It reveals a series of important and unresolved philosophical questions at the heart of Kantian ethics. It says that those of us who are in debt to Kant for so many key insights in moral theory will now have to build our own moral estimates of prenatal life at its various stages in independence from any stated positions of Kant's and on our own deepest understanding of his philosophy.



## 23. When Is a Person a Person – When Does the “Person” Begin?

*Natascha Gruber*

### 1. Introduction

One of the most polarized ethical disputes in human embryonic stem cell research (hESCR) is the question about the moral status of embryos: should embryos be treated as human beings, and as such, as potential persons, or as mere biological cell material, appropriate to be used for research? Since embryonic stem cells have the ability to differentiate into all types of cells of the human body, human embryonic stem cell therapy offers a hope for cure for severe diseases such as cancer, Alzheimer's, leukemia, or multiple sclerosis, among others. Despite its great therapeutic promise, hESC research is facing strong opposition since these stem cells can only be gained through the destruction of early stage embryos.

On the opponents' side of human stem cell research, arguments are often drawn from philosophy, to critique innovative technologies in life sciences, such as gene manipulation and cloning, and to support the ethical agenda to call for a restriction of hESCR. The main ethical references drawn from philosophy are concepts of personhood and human dignity, and one of the main arguments employed by hESCR opponents stresses that Kant's categorical imperative prohibits treating a person just as a means to an end (*Mittel zum Zweck*). Rather, persons are to be treated as ends in themselves (*Zwecke an sich*). For opponents, the very fact that embryos, from the moment of conception on, bear in themselves the potential of personhood, gives reason to call for an end, or at least for a significant restriction, of hESCR. Following this approach, human embryos should not be disaggregated to obtain stem cells for research and cloning, since, when human embryos are persons *in potentia* in a Kantian sense, destroying them to obtain their cells for research fails to treat them as ends in themselves. Consequently, opponents of hESCR claim that on the basis of their ontological status (as potential

persons), the same constraints that hold for killing adult humans apply to human embryos as well: killing cannot be justified with therapeutic promises of higher social ends. But the assumptions about the ontological as well as the moral status of early stage embryos in their first five days—so called blastocysts, and only these are used for harvesting stem cell lines—are controversially disputed in the current discourse, with the main question on the table being: when does a human being come into existence—when does the “person” begin? This question remains theoretically as well as empirically unsolved.

In this essay I will point out that if one wants to argue against hESCR by appealing to the Kantian concept of “personhood”, one has to bear in mind that this concept is rich and complex, since it presupposes the capacities of reason, free will, and moral agency. Only by possessing these features does a being, in the Kantian sense (human or not), have dignity, and herewith deserve respect and protection. Since in the current controversy, the line regarding the ontological and moral status of embryos cannot be drawn, not just opponents are using Kantian ethics and concepts to support their agenda, but proponents of hESCR are also able to draw on Kant to argue in *their* favor. Proponents question these assumptions of the “potential person” in an embryo and deliver quite challenging readings of Kant’s conception of personhood.<sup>1</sup> Manninen’s argumentation, for example—I will come back to it later in the essay—differentiates between biological and ontological categories, denying any causal relation between them. The main argument goes as follows: when Kant claims that humans have to be treated as ends in themselves, does that mean all members of the biological species *homo sapiens* have to be treated as ends in themselves? If so, then Kant would regard personhood as equivalent with being part of the biological species. But this is not the case, since Kant sets the very distinct definition: only intelligent beings (*vernunftbegabte Wesen*) are persons! Manninen’s argument draws a line between the biological and ontological dimension of the *homo sapiens* species and claims that, while a human being in the ontological sense (a person) is always also representing its biological species, the same does *not* hold vice versa: not every member of the *homo sapiens* species is a human being, such as an early stage embryo that biologically belongs to the *homo sapiens* species,

---

1 Bertha Alvarez Manninen, “Are Human Embryos Kantian Persons?: Kantian Considerations in Favor of Embryonic Stem Cell Research”, *Philosophy, Ethics, and Humanities in Medicine* 3.4 (2008), 1–16.

but is not regarded as a human being in the ontological sense, since it does not display the capacities of reason and moral agency. This position also denies that the event of conception would already causally entail the genesis of a person, since conception, cell fusion, and embryonic development are regarded as mere biological phenomena. This approach is clearly based on a distinction between biological/empirical phenomena on the one side and ontological concepts of human dignity and rational personhood on the other side. Consequently, the biological tissue of an early stage embryo does not constitute a person, and hESC research is justified.

These hESC cultures are derived from a blastocyst or early stage embryo (four to five days old), consisting of 50 to 150 cells. These cells are pluripotent, meaning they can develop into any of the more than 200 cell types of an adult body. Hence cultivated stem cells can be specialized to grow into various organs or tissues such as muscles or nerves. Whereas the medical application of embryonic stem cells is still in the state of basic research, with their therapeutic efficiency and applicability on adult patients not yet proven, adult stem cells, mostly gained from bone marrow, are already routinely used in medical treatments today.

Opponents disapprove of the scientific use *only* of embryonic stem cells while approving all other possible types of stem cell research. Therefore, much effort has been made in recent years to find methods for producing pluripotent stem cells, so called “induced pluripotent stem cells” (iPSC) artificially. These cells are derived via reprogramming of non-pluripotent adult cells, such as skin tissue and are regarded to possess the same, much desired capacity for differentiation as natural pluripotent stem cells, such as embryonic stem cells, do. This important achievement could in fact allow research with pluripotent stem cells *without* the controversial use of embryonic stem cells. Currently, scientific research on hESC as well as on iPSC is conducted, since it is not clear yet, whether iPSCs really do have the same qualities and therapeutic potentials of hESC cells; at this point, *neither* iPSCs *nor* hESC cells have been used on patients. Both lines are in the stage of basic, foundational research, and it may take at least another decade until firm results can be expected.<sup>2</sup>

Proponents, however, make yet another valid and in fact quite utilitarian point, why research should not abandon work with embryonic

---

2 J. Yu, et al., “Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells”, *Science* 318 (2007), 1917–20.

stem cells, despite the many other options.<sup>3</sup> The embryos that stem cell lines are drawn from are actually handed over to researchers as leftovers from fertility clinics, where thousands of abundant blastocysts are produced, frozen, and stored; Manninen mentions a number of about half a million within the US.<sup>4</sup> These frozen and stored blastocysts will be discarded and washed down the drain eventually. Since there is no way that these abundant blastocysts would ever be transplanted into a person’s womb, and brought to birth, they will die anyway, and, so the reasoning goes, why not use them for research that could serve and benefit all mankind?

What would Kant say if he lived today? Kant could not foresee the developments in technology the twentieth century has taken and the twenty-first century is heading into. So why refer to Kant? Isn’t that a highly speculative and scholastic enterprise? An abundance of material on Kant’s concepts of personhood and human dignity is facing scarce textual references Kant provided on the status of children, not to speak of the unborn. It takes a lot of exegetical analysis, as well as an in-depth overall understanding of Kant’s philosophy and ethics, to answer the speculative question: which moral status would Kant ascribe to embryos? But no matter how one draws on Kant, his relevance for today’s as well as for future discourses is unbroken as ever, since his ethics is still the only normative reference Ethics Commissions can draw on as an alternative to utilitarian, pragmatic, and other types of reasoning.

In the following parts I will give an overview of the main dispute in stem cell research, then take a closer look at Kantian arguments, referring to the *Groundwork of the Metaphysics of Morals*, the *Metaphysics of Morals*, and the *Critique of Pure Reason*. In the summary and concluding part I will argue against those who claim that Kant’s ethics is compatible with hESC research, or that the justification of hESC research is *derivable* from this ethics. I doubt this approach, and I will try to support my point through a textual exegesis, although I have to admit that the references are quite scarce. Coming to the conclusion that a justification of embryonic stem cell research is not derivable from Kantian ethics does not mean I personally oppose this research. But from my point of view as a Kant scholar I will argue that *if* someone wants to provide argu-

---

3 Katrien Devolder, “Human Embryonic Stem Cell Research: Why the Discarded-Created Distinction cannot be Based on the Potentiality Argument”, *Bioethics* 19 (2005), 167–86.

4 Manninen, 13.

ments for embryonic stem cell research, one would not find support for them within Kant's philosophy.

## 2. The Question of the Ethical Status of Human Embryos in the Current Debate

Why is it that the moral status of human embryos is so controversially discussed and that this question is able to polarize the discourse so strongly? Maybe it is because this question entails one of the most fundamental issues regarding the nature of human life. To be able to draw the line on what is human life and what is not seems to be crucial, since the status of a human being commands dignity, respect, and the right to live.

So, is a five-day-old human embryo a human being or not? While this question is in dispute, it is *not* in dispute that a five-day-old embryo is a *living being*. For this fact, hESC research opponents, including Christian and other religious and humanitarian groups, claim the classic view that, since human life begins to exist at the event of conception, when an egg and sperm fuse to form a one-cell zygote, human embryos should already at this early stage be regarded as living members of the human society with the *potential* to become adult persons. As such, they deserve protection and the right to live. On the opponent's argument against hESC research, membership in the *homo sapiens* species confers on the embryo a right not to be killed. This view is grounded in the assumption that human beings have the same moral status at all stages of their lives, as soon as they come into existence as a living entity.

Proponents for hESC however have, as mentioned before, developed elaborate arguments to reject giving human embryos the status of human beings. Apart from the biological/ontological distinction, the point is made that the cells of blastocysts do not in any way form a human organism, since these cells are not differentiated but rather homogeneous. Cells start to grow into a human embryo after cell differentiation, usually starting from day 14 to 16. Although the cells of blastocysts are in fact living cells, they are *not* regarded as a human organism; so again, research with them is justified.<sup>5</sup> From this approach the conclusion is drawn that species membership, as is undoubtedly the case

---

5 J. McMahan, "Killing Embryos for Stem Cell Research", *Metaphilosophy* 38 (2007), 170–89.

with human zygotes and blastocysts, does not entail the ontological or moral status of a human being, with a right to live. Instead, higher order capacities, such as reasoning, self-awareness, and moral agency are claimed as criteria to ascribe a being the right to life. But basing the foundation for a right to live in capacities such as reasoning, self-awareness, and moral agency entails the difficulty that human newborns and infants lack these capacities, even to a greater degree than some non-human animals, as chimpanzees, do. To challenge this difficulty, the potentiality argument comes in again, with the attempt to conceptualize a distinction between “exercisable capacities” and “basic natural capacities” as innate, inborn presumptions for higher mental capacities. Exercisable capacities are current actualizations of these innate natural capacities.<sup>6</sup> Following this approach these basic natural capacities exist already in an early stage of embryonic life. The difference between these types of capacities is regarded as a difference between certain *degrees of actualization* along a developmental continuum line. In fact there are differences in actualization between the capacities of embryos, fetuses, infants, children, and adults.<sup>7</sup>

But the question is: do these differences of actualization justify the introduction of the same moral and ontological standards for all of an individual’s developmental stages? Again, proponents deny that being endowed with a certain potentiality would logically entail the same status as having *realized* some or all of these potentials.<sup>8</sup> Furthermore, so the argument goes, if the basis for protecting embryos were grounded in their potentiality to grow into human, intelligent beings, the thousands of frozen and stored cells, in order to realize their potential, would need to be implanted into (willing) females’ wombs—an idea that would raise

---

6 R. P. George and A. Gomez-Lobo, “Statement of Professor George and Dr. Gomez-Lobo”, in: *Human Cloning and Human Dignity: An Ethical Inquiry. President’s Council on Bioethics* (Washington: Council Publications, 2002), 258–66. Available online at: [www.bioethics.gov](http://www.bioethics.gov).

7 Allen Wood, in his reading of Kantian ethics, draws a distinction between persons “in the strict sense” and persons “in the extended sense”. Whereas persons in the strict sense possess the full range of capacity for reason and moral agency, individuals in the extended sense (including children) would expose only partial stages of rationality, or preconditions of it. See Allen Wood and Onora O’Neill, “Kant on Duties Regarding Nonrational Nature”, *Proceedings of the Aristotelian Society*, Supl., 72.1 (1998), 189–228.

8 A. Sagan and P. Singer, “The Moral Status of Stem Cells”, *Metaphilosophy* 38 (2007), 264–84.

serious concerns about the dignity and moral status women possess.<sup>9</sup> These problems raise serious doubts regarding an embryo's potential and the potentiality argument as a foundational argument for a right to live, suggesting the conclusion that the moral status of early stage embryos is not great enough to restrict research that may yield valuable therapeutic benefits for all mankind.

### 3. A Kantian Analysis of Embryonic Stem Cell Research: Potentials and Limitations

Kant did not deal with the question, so crucial for the current debate: when is a human a human, a person a person? Kant's starting point is: humans are intelligent beings (*vernunftbegabte Wesen*) having intrinsic moral value and moral dignity. As such they are ends in themselves, and not just means to an end. In the *Groundwork of the Metaphysics of Morals* Kant sets the axiom: intelligent beings exist as ends in themselves (als *Zwecke an sich*). From this axiom the categorical imperative, in its four different formulations, is derived; I pick here the second, the practical imperative toward mankind:

... Now I say: that the human being and in general every rational being exists as an end in itself, *not merely as a means* for arbitrary use by this or that will: ... instead he must always be regarded at the *same time as an end* ... The practical imperative will therefore be as follows: Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end.<sup>10</sup>

Kant draws a clear distinction between persons and things. Whereas persons have the capacity to reason, things (including animals) do not have this capacity. According to Kant, only beings with the capacity to reason and the capacity for moral agency are subject to dignity and respect. Be-

---

9 Susan Feldman, "From Occupied Bodies to Pregnant Persons: How Kantian Ethics should Treat Pregnancy and Abortion", in J. E. Kneller and Sydney Axinn (eds.), *Autonomy and Community: Readings in Contemporary Kantian Social Philosophy* (Albany: New York State University Press, 1998), 265–82; Ronald M. Green, "Is There a Kantian Perspective on Human Embryonic Stem Cells?", in Stephen Palmquist (ed.), *Cultivating Personhood: Kant and Asian Philosophy* (Berlin: Walter de Gruyter, 2010), ch.22.

10 Immanuel Kant, *Groundwork of the Metaphysics of Morals*, tr. H. J. Paton (New York: Harper Torchbooks, 1964), 95–6.

ings without the capacity to reason (i. e., “things”) do not have dignity; this entails that we do not owe them the same kind of respect we owe beings with the capacity to reason. On this basis, the argument goes on further and links with a being’s capacity for moral agency, as the capacity to set and perform moral laws within the community of moral beings, the “realm of ends”. By the “realm of ends” Kant denotes the community of intelligent beings, regulated through moral laws all intelligent beings are subject to, and where no one regards her/himself and other fellow members just as means, but rather as ends in themselves. Within the Kantian framework, the capacity for free will and moral agency *actually* entails dignity, due to the fact that intelligent beings are *at the same time* also moral beings and, as such, legislative as well as subject to moral laws.

Another interesting differentiation Kant makes is the one between “price” and “dignity”: in the realm of ends, he says, everything has a price *or* a dignity. What has a price is replaceable; what is priceless, and is, as such, above all pricing, has dignity (4:102). This distinction obviously follows from the thing/person distinction, and it is easy to see what goes with what. Human embryos, produced in fertility clinics, surely have a price; in fact, a quite high one, but do they also have dignity? In the exclusive reading of the “or”, what has a price does not have dignity, and vice versa: what has dignity, does not have a price. Since this distinction in Kant’s *Groundwork* suggests that the price/dignity distinction goes along with the thing/person distinction, can we now logically conclude that, according to Kant, *since* human embryonic stem cells *do* have a price, they do not have dignity?

Due to the fact that Kant leaves so much open, both opponents as well as proponents of hESC research are able to draw on Kantian ethics in their favor. If within the Kantian framework stem cells can be regarded as “things”, they only have a relative value and can be used as means, for higher ends, and therefore research is justified. If they are regarded as rational beings (*in potentia*), they would have an absolute value and their consumption for research would not be justified.

In order to support the being or person *in potentia* argument, one has to browse Kantian texts other than the *Groundwork* or the *Critique*, since, as mentioned, Kant did not incorporate children and the unborn into his philosophy. To my knowledge, the only statement on the moral status of children can be found in the *Metaphysics of Morals* (Rechtslehre, §28); due to its exegetical value, let me quote the whole length of the passage (6:280–1):



... children, as persons, have by their procreation an original innate (not acquired) right to the care of their parents, until they are able to look after themselves ... for the offspring is a person, and it is impossible to form a concept of the production of a being endowed with freedom through a physical operation, so from a *practical* point of view it is quite a correct and even necessary idea to regard the act of procreation as one by which we have brought a person into the world without his consent and on our own initiative, for which deed the parents incur an obligation to make the child content with his condition so far as they can. They cannot destroy their child, as if it were something they had *made* (since a being endowed with freedom cannot be a product of this kind) or as if it were their property; nor can they even just abandon him to chance, since they have brought not merely a worldly being but a citizen of the world into a condition which cannot now be indifferent to them even just according to the concepts of Right.

With the act of procreation a person is brought into being, is “drawn into this world” without his/her approval, says Kant. Furthermore, Kant makes the interesting and for the current interpretation crucial point that it is impossible, to obtain an idea about how a being, endowed with reason and freedom, is brought into this world via “physical operation” (here: conception), giving a hint to the puzzle of how a biological event is able to render beings endowed with reason and freedom.

This is the vague line the Arizona scholar Bertha Alvarez Manninen draws on in her paper “Are human embryos Kantian persons?” Manninen provides a very appealing and challenging interpretation to create strikingly supportive arguments for hESCR, based on Kantian philosophy. With Kant she reads that it is impossible to understand how beings, endowed with reason and freedom, come into existence through the physical occurrence of conception.<sup>11</sup> Furthermore, Manninen draws on the first *Critique*, where Kant elaborates the tension that humans are biological as well as intelligent beings. As such, they are subject to natural laws, hence causally determined, as well as being not causally determined, but free. This tension Kant tries to resolve with the phenomenal (empirical) and noumenal (intelligible) distinction. Indeed, these two spheres are crucial within Kant’s theoretical framework, since the first does not causally influence the second. So, to claim the biological event of conception *causally* entails the existence of a being, endowed with reason and freedom, seems to contradict Kant’s statement that we cannot understand the creation of a free being from a purely physical

---

11 Manninen, 8.

operation. Kant’s conclusion, formulated in the Third Antinomy, is that we cannot ground the existence of transcendental freedom by referring to the phenomenal world (A448/B476 f).

I agree with Manninen’s argument; but the noumenal/phenomenal distinction bears unsolved tensions in Kant’s philosophy, so the interpretation can also go the other way round: *if* empirical (“physical”) operations cannot causally render the creation of a free being, then *nowhere* along the line of the development of a human being can the coming into “existence” of reason and freedom be pointed out, because these developmental stages from zygote, blastocyst, embryo, infant, and so on, are *all* empirical or biological, phenomena. So this leaves open when the noumenal or intelligible *causally* comes into play, unless it has already always been there! But that brings us back again to the potentiality argument, since Kant seems to indicate that the noumenal ground has been there all along.

It may be worth taking a closer look at the wording in the cited passage above: whereas the English translation uses “bringing a person into the world”, Kant’s German uses the verb “*herübergezogen*”, literally meaning “drawn to ... from ...”; this is actually quite a strong verb, indicating an operation of movement, of a drawing, or pulling of a subject from point A to B. Why would have Kant chosen exactly this particular verb? In the overall framework of Kant’s distinction between the empirical and transcendental, this wording could suggest that *via* the act of procreation a being is drawn *from* the “intelligible” *into* the phenomenal world. Given the cultural and historical context of Kant’s time, regarding prevalent views of the immateriality and immortality of the soul that Kant was also dealing with, I think this is a possible reading of what Kant might have been indicating.

The incorporation of Kantian ethics into bioethical discourses still leaves open the question: where do the capacities of reason, self-awareness, and freedom come from? The interpretation of Peter Baumanns also denies that supportive argumentation for hESC and other controversial bio-technologies can be drawn from Kant’s ethics. Baumanns states “that the personal and moral status of the human embryo is one of the philosophically unresolved problems of the current bioethical discussion”<sup>12</sup> and finds that Kant’s concepts of the individual as autonomous, within the community of the autonomous, entail the idea of

---

12 Peter Baumanns, *Kant und die Bioethik* (Würzburg: Königshausen & Neumann, 2004), 5.

an embryo as a moral subject to come. Baumann introduces into the discussion the term *nasciturus* (from Latin: “one who is to be born”) to indicate the embryo’s special status, not just as plain biological cell material but, in line with Kant, as a world being and world citizen about to become. For Baumanns the “drawing into being” of a “person in potentia” into this world, into the community of humans, via conception, pregnancy, birth, and education, is not reducible to biological processes, and therefore all developmental stages of human lives deserve protection.

#### 4. Summary and Conclusion

As modern reproduction technology with its high tech equipment like ultrasound, x-rays, etc.—and Kant had seen none of that in his days—demonstrates so impressively, every moment of procreation and pregnancy can be traced and observed but also manipulated and altered. But no matter how severely we ultrasound uteruses, no matter how thoroughly we scan human brains, all we can “see” is still plain cell material, all we can “observe” are biophysical, biochemical events and phenomena—“physical operations” as Kant would have put it. No neuroscientist has ever been able to locate the “I”, a free will, and this not due to the poor level of technology, with the hope for future generations to detect. With Kantian philosophy it can be shown that no neuroscientist will ever be able to “see” an “I”, a “free will”, since these are not empirical phenomena, and as such remain the unsolved puzzles of human existence. “What it means to be human” remains invisible, immeasurable, and unobservable at its beginning—wherever this beginning is, or comes into play: at the event of conception, in the first weeks, in the ninth month, in an infant’s first years, or in an adolescent teenager. Kant could not unveil this puzzle, nor can we, despite the high-tech equipment available today.

Digging into the bioethical discussion around stem cell research gave me the impression that what goes on is a highly scholastic dispute between pro and con; instead of adding to this, I would like to take a step back, to get a look at the bigger picture. Let us draw attention away from the question of the moral, metaphysical, or ontological status of a zygote, blastocyst, or embryo, and reflect upon the reproduction industries that produce these entities in abundance, as we have learned. Financially potential couples and single women invest thousands of dol-

lars for a genetically bio-child, but why, when there are plenty of babies and children waiting for adoption? Where does the “desire” to have one’s own bio-kids come from; through which cultural practices and values is this desire established and nurtured, allowing reproductive industries to capitalize and earn millions of dollars? These issues I found nowhere addressed in the current discourse.

As the development in biotechnology progresses, in the end the potential medical benefits gained from hESC research will probably outweigh the loss of embryos involved. As I stated above, I do not oppose embryonic stem cell research, but I would not support my argument with Kantian ethics. Zygotes and blastocyst cells are artificially produced in tubes. From a certain developmental stage on, outside a uterus, a woman’s womb, these entities could not exist. Therefore I agree with the statement of Rabbi Elliot Dorff, cited in Manninen’s paper, who finds that “extracorporeal (ex utero) embryos have no legal or moral status outside the womb under Jewish law because, “outside the womb ... they have no such potential to become persons”.<sup>13</sup>

Another rather complex relationship has come to the surface through the bioethical discourse: the relationship between technology, science, and society, the dialectics between scientific progress versus ethical values and standards, calling for certain restrictions, and the need for ethical norms. The search and quest in life sciences will go on—no doubt about that—and will always try not just to go to its limits but to transgress them. This principle, that what drives science—the unsolved (and unsolvable) questions and search for answers—will never come to an end, we can already find in Kant’s first *Critique*; see, for example, the First Antinomy (A426/B454 f). Are there any boundaries that should not *ever* be transgressed? I have to leave this to the Science and Ethics Commissions.

---

13 Elliot Dorff, “Testimony for the National Bioethics Advisory Commission”, in M. Ruse (ed.), *Stem Cell Controversy: Debating the Issues* (Amherst, NY: Prometheus Books, 2003), 197, quoted in Manninen, 16.

## 24. Personhood and Assisted Death<sup>1</sup>

*Anita Ho*

### 1. Introduction

Many have argued that Kant's direct statements on the prohibition of suicide imply that medically-assisted deaths are impermissible, since they destroy one's personhood and contradictorily will to end a moral being's reasoning. Exploring various texts, this essay argues that Kant's writings do not clearly prohibit all hastened deaths. While Kant's discussions suggest that rational agents who are experiencing irreversible physical decline cannot will the destruction of their personhood and commit suicide or request assisted deaths, non-voluntary or even involuntary euthanasia<sup>2</sup> may be allowed in situations where patients have lost their rational capacity, regardless of their physical condition. Through an exploration of Confucian understandings of human life, this essay ends by suggesting that attention to relational aspects of personhood may be necessary to help explain why non-voluntary and involuntary euthanasia are morally problematic.

### 2. How Kant Explained the Moral Problems of Suicide

In his various writings Kant explicitly says no moral agent can consistently and rationally will the generalized form of any suicide maxim based on self-love as a law of nature; such generalization will result in a contradiction because self-love's function is "to stimulate the furtherance of life", not to promote its end (4:89). Like the sages of ancient

---

1 For constructive comments on earlier drafts, I thank Stephen Palmquist, Suze Berkhout, and participants at the Kant in Asia International Conference (2009) and the International Conference on Applied Ethics (2009), respectively.

2 This essay defines euthanasia as the act of painlessly ending the life, such as by lethal injection, of an individual experiencing irreversible decline from a terminal illness or an incurable condition.

China, who thought it was human being's inborn nature to cherish life and be delighted in survival,<sup>3</sup> Kant believes self-preservation is part of the laws of nature. Suicide cannot be based on the exercise of one's free will, since by ending one's life the person is eliminating a necessary feature of moral agency. Such action puts oneself "below the beasts" and is incompatible with the formula of humanity, evoking "revulsion with horror" (27:372) and making it "impermissible and abhorrent" (27:375). A person who commits suicide when a longer life "threatens more evil than satisfaction" is using oneself "as a means to maintain a tolerable condition up to the end of life" (6:422). Physical pain does not diminish the value of a person but only the worth or desirability of one's condition—even though the moral person and the person's physical body are connected, they are not identical.<sup>4</sup> Arguments for suicide based on physical sufferings only mistake the person's phenomenal condition for the noumenal self.<sup>5</sup> Rational persons, whose desire to prevent physical pain and suffering is understandable and consistent with "the rule of prudence" (27:373) nonetheless may not use any means that would contradict "the rule of morality" or their natural duty of self-preservation. Even if a longer life threatens more evil than satisfaction, disposing of oneself as a mere means to the discretionary end of terminating suffering debases humanity in one's person; such "murdering [of] oneself" not only annihilates the subject of morality in one's own person; it roots out the existence of morality itself from the world (6:422).

Based on this line of argument and his explicit statement that "suicide is not permitted under any condition" (27:372) it seems Kant would reject any autonomy-based argument for voluntary euthanasia or assisted deaths, despite one's physical suffering at end-of-life. However, in recent years, some have suggested that Kant's statements regarding the danger of annihilating humanity do not preclude all forms of (assisted) suicide at end of life. Two arguments are particularly salient. First, it has been argued that altruistic suicide is compatible with the formula of humanity, because the suicide is not motivated by self-serving

- 
- 3 M. Shen, "To have a Good Birth as Well as a Good Death: The Chinese Traditional View of Life and its Implications", in R. Qui (ed.), *Bioethics: Asian Perspectives* (Boston: Kluwer Academic Publishers, 2004), 135–46; esp. 138.
  - 4 C. Perry, "Suicide Fails to Pass the Categorical Imperative", *American Journal of Bioethics* 7.6 (2007), 51–3.
  - 5 H. L. Nelson, "Death with Kantian Dignity", *Journal of Clinical Ethics* 7.3 (1996), 215–21; esp. 217.

ends and hence the person is not treating oneself merely as a means to an end.<sup>6</sup> Second, some have argued that, even if rational persons may not end their life for physical suffering, Kant's focus on preserving moral agency may allow or even require those who are facing the loss of personhood to commit suicide or seek assisted deaths.<sup>7</sup>

### 3. "Altruistic" Suicide at End of Life

Consider the former argument, that not all forms of suicide treat one as a mere means. Altruistic suicide, for example, is not motivated by personal suffering or self-serving ends but by one's concern for burdening another.<sup>8</sup> As technologies continue to extend people's lives, many now live with chronic conditions and various disabilities in old age and require substantial caregiving from family members or others. Respect for our loved ones demands that we honor their freedom to exercise their own moral agency. If by extending her life an elderly mother is using her family as a means of contributing to her maintenance despite the cost of great misery to them, she is considering the intimate relationship as simply instrumental and treating the loved ones "as means to her own ends."<sup>9</sup> When a suicide is motivated by one's beneficent concern for another rather than a subjective desire to relieve one's own pain, it is allegedly consistent with treating oneself as an end.

Certainly, manipulating or forcing others to care for us regardless of their own desires and ability to do so would violate both our perfect and imperfect duties toward them.<sup>10</sup> In respecting the family members' autonomy or moral agency, a patient has a perfect duty to refrain from *imposing* caregiving work on others. Universalization of such imposition

---

6 Nelson, "Death with Kantian Dignity", 218.

7 D. Cooley, "A Kantian Moral Duty for the Soon-to-be Demented to Commit Suicide", *American Journal of Bioethics* 7.6 (2007), 37–44.

8 Nelson, "Death with Kantian Dignity", 218.

9 Nelson, "Death with Kantian Dignity", 219.

10 4:421–32. According to Kant, perfect duties admit of no exception in the interests of inclination. They require adherence without exception and specify a particular action (e. g., do not lie). We have a perfect duty not to act by maxims that, by universalization, would result in logical contradictions. Imperfect duties such as beneficence to others demand the adoption of an end or certain maxims rather than specific actions. They may be overridden by other imperfect and perfect duties and allow a significant degree of freedom in deciding how to comply with them.

would treat humanity as a mere means and refute the meaning of autonomous decisions. Moreover, one has an imperfect duty of beneficence toward others' well being, including that of one's own family. Even though such imperfect duty does not "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" (6:390), what David Cummiskey calls a robust rather than an anemic interpretation of imperfect duties would clearly forbid us to adopt a maxim of indifference to the welfare of those who are caring for us.<sup>11</sup> While interdependency is part of the human condition, and thus it is unclear whether caregiving work should be considered a "burden", such work can be financially, physically, and emotionally taxing for care providers,<sup>12</sup> making it important for patients to keep in mind the well being of these family caregivers in their end-of-life planning. As Confucius also states, benevolence is more important than life—we may not violate benevolence for the purpose of survival.<sup>13</sup>

Nonetheless, even considering one's perfect and imperfect duties toward the caregivers, "altruistic" suicides are inconsistent with Kant's moral framework. Imperfect duties can be limited by other imperfect duties and trumped by perfect duties, and Kant's hierarchical ordering of these two types of duties would seem to require that any altruistic intention to spare the loved ones burden must be considered in the context of one's perfect duty of self-preservation. While time and resources used to care for a frail and declining patient can be progressively substantial, their subjective disvalue is limited and superseded by the objective and unlimited value of personhood. Since suicide, even if done out of beneficence toward our loved ones, still contradicts our perfect duty to preserve our lives, we need to find other options in promoting our caregivers' well-being and refraining from imposing unreasonable sacrifices on them.

Such alternatives may be sought at the societal level. For example, the society can consider providing services, programs, and funding to make sure family members, particularly girls and women, are not disproportionately burdened with caregiving work. As Kant notes in his dis-

---

11 D. Cummiskey, *Kantian Consequentialism* (New York: Oxford University Press, 1996), 109–37.

12 Notice that family members may also have an imperfect duty of beneficence to care for the patient at end of life. Such a duty is not limitless; it is constrained by the family members' ability to meet other imperfect and perfect duties.

13 Shen, "To have a Good Birth as Well as a Good Death", 139.



discussion on our duty to help those struggling with great hardships, the possibility that one “would need the love and sympathy of others” renders a maxim of indifference to contributing to others’ welfare in time of need contradictory (4:423). Practical reason requires that we uphold *with all others* the duty of mutual benevolence according to our means, and make the happiness of both ourselves and others our end (6:450–4). Recognition of our interdependency can help to design a social system that would facilitate our mutual duties of benevolence by considering caregiving work as communal responsibilities rather than individual cases of burden.<sup>14</sup> Conceptualizing caregiving responsibility as more than individual work may also help to ease the persistent inequity in the distribution of family and paid care work between genders.

The notion of altruistic suicide is also conceptually questionable within Kant’s framework of what constitutes a moral action. First, while the adjective “altruistic” supposedly denotes a moral form of suicide, killing oneself on the alleged basis of not burdening others would at best be non-moral based on Kant’s criteria for moral actions. An action has moral worth only if it is done *from* duty (i. e., purely motivated by one’s respect for the moral law or “internal lawgiving”) (6:220, 392). Many people who seek (assisted) suicide likely are motivated not only by their duty of beneficence but also by a desire to avoid further pain and suffering. While certain desires or inclinations may facilitate the effectiveness of and conform to our moral duty of beneficence in some situations, an action that is moved by a personal interest is nonetheless heteronomous, rendering it impossible to ascertain its moral worth. In fact, we can see how such mixed motives can lead one astray. If one’s inclination to ease one’s suffering may tempt that person to ignore the perfect duty of self-preservation, even the additional “altruistic” attitude would not save such suicides from being immoral under Kant’s framework.

Second, without romanticizing caregiving work, it is noteworthy that many appreciate the opportunity to care for a dying relative. And in some cultures, including the Chinese culture, it is not unusual for an elderly sick person to be dependent on his/her children as something

---

14 J. Gentzler, “What is a Death with Dignity?”, *Journal of Medicine and Philosophy* 284 (2003), 461–87.

s/he is entitled to.<sup>15</sup> According to the Confucian teaching, children have a fundamental and comprehensive responsibility to care for their parents as part of their moral cultivation.<sup>16</sup> Those who adopt the Confucian philosophy may also consider exercising their filial responsibility as strengthening emotional connection to their aged parents and part of living a virtuous life. They may not see each other as separate individuals, but rather as members of a coherent unit. As I have also argued elsewhere, some people's identities are rooted in their social and familial affinities, making the relationship between caregivers and patients interdependent and more complex than often acknowledged.<sup>17</sup> Despite substantial sacrifice or inconvenience, many caregivers willingly take on such responsibilities as an expression of their love and a precious chance to maximize the time they have with the person. The legitimate feminist concern of women being expected to carry out caregiving work notwithstanding, many patients who are close to their families are unlikely to be using their intimates as *mere* means—the former generally love and respect their caregivers, appreciate what they are doing, and worry about how such caregiving work may affect their own well-being. Certainly, as mentioned earlier, individuals should be allowed to decide in accordance with their various duties and goals, and patients should not pressure people to take on caregiving work beyond their capacities, especially given how such work has historically been imposed on women as a social group. Nonetheless, altruistic suicide (regardless of family members' caregiving desires) may paternalistically override these moral agents' autonomy. In cases where a patient seeks suicide because the family is not willing to take on caregiving work, such pressure would render the suicide involuntary rather than altruistic; this would likely concern Kant, who says nobody can oblige us to commit suicide (27:371).

---

15 E. Hui, "Personhood and Bioethics: Chinese Perspective", in R. Qui (ed.), *Bioethics: Asian Perspectives* (Boston: Kluwer Academic Publishers, 2004), 29–44: esp. 41.

16 R. Fan, "Which Care? Whose Responsibility? And Why Family? A Confucian Account of Long-Term Care for the Elderly", *Journal of Medicine and Philosophy* 32 (2007), 495–517.

17 A. Ho, "Relational Autonomy or Undue Pressure? Family's Role in Medical Decision-Making", *Scandinavian Journal of Caring Sciences* 22 (2008), 128–35.

#### 4. Suicide When Facing Loss of Personhood

Recently, some have argued that, even if Kant's ethical framework may not allow rational persons to end their life for physical pain and suffering, his focus on preserving moral agency over physical life may allow or even obligate those who are facing imminent loss of personhood to commit suicide. Physical life is insufficient for moral agency, and in situations where agents are faced with the alternative of keeping either their physical or moral life but not both, sacrificing their body to prevent them from degenerating into a thing that cannot act from the autocracy of the human mind may be more important.<sup>18</sup> Dementia, for example, is an irreversible condition that progressively impairs many elderly persons' cognitive function. It limits their memory, language and problem-solving skills, attention span, and ability to make rational decisions or to understand their situations, ultimately taking away their personhood. Due to this lack of moral agency and potential to harm others, these individuals' freedom can sometimes be constrained against their wishes without violating their autonomy. In various jurisdictions, individuals with moderate to severe dementia are sometimes certified and treated involuntarily when the police, a judge, or a physician determines they may cause significant disruption and harm to their or others' lives. While similar actions against rational patients would be considered paternalistic or even severe violations of their dignity and autonomy, they are sometimes deemed morally (and legally) acceptable for individuals who do not possess full selfhood.<sup>19</sup>

The moral significance of personhood in distinguishing what we can and cannot do to various human beings in Kant's philosophy seems to call for different considerations regarding suicide and assisted deaths for people with limited cognitive functioning. While rational persons who take their own lives or seek assisted death to avoid suffering commit a moral wrong because they are treating themselves as mere means, such an argument leaves open the question of whether suicide for those who are facing progressive loss of personhood, and hastened deaths for patients who are unconscious or no longer possess rational capacity, violate Kant's moral framework. Even though Kant clearly prohibits sui-

---

18 Cooley, "A Kantian Moral Duty ... to Commit Suicide", 37.

19 It is important to acknowledge our fallibility and potential bias in predicting the likelihood and extent of harm. However, a detailed discussion of these issues is beyond the scope of this essay.

cide in cases where one's mental capacity is intact, sacrificing physical life may be congruent with or even required by one's duty to preserve one's moral life and inner worth in cases where life extension is no longer in keeping with the dignity of humanity (27:377). This coincides with Confucius' teaching, which says that a person of humanity will never seek to live at the expense of injuring humanity. Such a person would rather sacrifice his or her life in order to realize humanity.<sup>20</sup> According to Kant, what makes us human is not our physical life that we share with all animals, but our unique moral nature. Our physical life, while supporting our moral life, is not a necessity; living honorably while life lasts is a necessity (27:373). In those rare cases involving a choice between physical and moral lives, a rational agent ought to preserve the latter. Citing Kant's discussion of Cato and Lucretia, who allegedly had "honourable motives" for ending their lives, and his casuistical question regarding a person who may lose his own humanity and "harm others as well in his madness as a result of being bitten by a rabid dog," (6:423–4), some argue that Kant would support suicide to preserve one's personhood.<sup>21</sup>

It is important to note that, even though Kant believes we should seek to sustain our lives only insofar as we are worthy to live, it is unclear that patients facing cognitive decline are unworthy to live, such that suicide or assisted death would be required to preserve their moral worth. Even if some conditions are undesired, the aforementioned distinction between one's phenomenal condition and noumenal self would caution against any assumption that those who are facing declining rationality are not worthy to live. As many disability scholars and activists have warned, stereotypical and discriminatory attitudes toward the worth of people with disabilities have been behind the eugenics movement and many social policies that continue to disadvantage and demean such people. Even though dementia is progressive, irreversible, and would inevitably lead to death, it is unclear that Kant would allow suicide or assisted death for patients facing such decline, given that Kant distinguishes between "a suicide and one who has lost his life to fate" (27:371). It is one thing to die of dementia, a condition one has no control over. It is another to kill oneself because one finds living in such condition unbearable. While Kant emphasizes the importance of preserving our dignity, he does not suggest that suicide is the only way

---

20 Hui, "Personhood and Bioethics", 39.

21 Cooley, "A Kantian Moral Duty ... to Commit Suicide".

to do so, even in the most challenging cases. Kant himself does not offer an answer to his question regarding whether the bitten man who might go mad and harm others should take his own life. Even for Cato, who killed himself to preserve the lives of many other Roman people upon learning that the Republic was lost to Caesar, and Lucretia, the wife of a Roman leader who sought death after being raped because she felt she could no longer express a purity of conscience to the world, Kant insists they should adhere to their resolve with steadfast mind and defend to the death their honor. Those with such responses, rather than laying hands upon themselves, would show real courage and nobility (27:371–5).

Most patients with dementia do not pose significant harm to others, and for those who may, there are other ways to minimize such harm.<sup>22</sup> Moreover, as some have pointed out, dementia is a gradual process that “ebbs and flows”, with “moments where personhood shines through.”<sup>23</sup> While suicide could prevent a person from going through the potentially agonizing and frightening process of cognitive decline, if a person commits suicide prior to *complete* loss of personhood, the suicide would destroy any remnants of his or her rational agency, making it premature and problematic from a Kantian perspective.<sup>24</sup> The idea of preserving one’s moral agency by ending one’s physical life thus appears contradictory, since it destroys the necessary physical conditions required for any rational action and moral deliberation. Even though Kant’s focus is on our moral life, he contends that “the body is the total condition of life” (27:369). Given that “we have no other concept of our existence save that mediated by our body, and since the use of our freedom is possible only through the body,” premature destruction of the body also eliminates the “person” and “the power of choosing itself” (27:369).

What about suicide or assisted death for individuals who no longer possess self-awareness or will never develop such consciousness?<sup>25</sup> After

---

22 Some of these other measures, such as restraining a patient, would also restrict the person’s freedom, raising questions of how we should balance the right of an individual and that of others who may be harmed.

23 Perry, “Suicide Fails to Pass the Categorical Imperative”, 52.

24 M. Gunderson, “A Kantian View of Suicide and End-of-Life Treatment”, *Journal of Social Philosophy* 35.2 (2004), 277–87; esp. 280.

25 For our purpose here, I will not distinguish the different reasons for these humans’ permanent loss of self-consciousness. My focus is only on the implications of the lack of personhood on a Kantian ethical framework of euthanasia.

all, if there would no longer be any freedom, the destruction of the body would not eliminate any personhood that has already ceased to exist. Since the notion of suicide presupposes the possibility of voluntary behavior,<sup>26</sup> those who are incapable of acting out of their own will cannot, by definition, commit suicide. Any form of assisted or self-inflicted death would be at most non-voluntary. Nonetheless, it is unclear whether Kant's argument against suicide among rational beings would preclude euthanization of individuals who are no longer persons. While a person possesses an absolute inner worth whereby s/he deserves respect from all other rational beings, Kant says little regarding what we cannot do to human beings who are not subjects of a morally practical reason, or whether they possess inviolable dignity. Kant thinks of non-human animals and the animal side of human life as belonging to the deterministic realm of nature.<sup>27</sup> Unlike persons, whom by virtue of their capacity for moral rationality rise above that realm and exist as well in the realm of ends, Kant alleges that even sentient animals are not self-conscious and thus have no independent value—they only have value in relation to human ends. Kant only worries that how we treat sentient animals, which share with us a similar tendency to reciprocate loyalty, may influence how we treat persons. Rejecting the idea that we have direct moral duties to animals, Kant relies on the presumed human psychology that when we are kind toward animals, we strengthen the disposition to behave similarly to persons.<sup>28</sup> Persons have indirect duties toward animals only in the sense that certain actions toward these beings may damage our natural sympathies and lead us to mistreat rational moral agents.

Analogously, if human beings who have permanently lost their reasoning capacity and thus their personhood are no longer moral agents that require respect, then it would seem that we have no direct duties toward them. Even if assisted suicide is a contradictory term for humans who are not self-conscious, non-voluntary or even involuntary euthanasia for suffering patients who have lost their rational capacities are possible and potentially legitimate. As long as such killings are not carried out in a cruel manner, however defined, and that there are ways to en-

---

26 Gentzler, "What is a Death with Dignity?", 468.

27 M. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, Mass.: The Belknap Press of Harvard University Press, 2006), 131.

28 Nussbaum, *Frontiers of Justice*, 330.

sure that hastening these humans' deaths would not harden our dealings with self-conscious patients, euthanasia for humans who have permanently lost their rational capacity would not violate humanity. In fact, cruel killings of these "non persons" would not be problematic in themselves; they would be so if and only if they may influence our treatment of self-conscious persons. Hastening the deaths of patients who are no longer persons may help promote the priorities for other persons, such as by being beneficent toward those who may otherwise have to take on substantial caregiving work or by saving resources that can be reallocated for self-conscious patients.

### 5. Conclusion: Can Asian Relational Notions of Personhood Help?

Opponents of physician-assisted suicide and voluntary euthanasia often argue that abuse may happen: some who inquire about hastened death may be pressured into doing so. Defenders who acknowledge such concerns regularly respond by proposing safeguards to ensure that only those who are competent and not under other pressures can request hastened deaths. Any Kantian argument for euthanizing those who have purportedly lost their personhood while denying voluntary euthanasia and assisted suicide would therefore likely raise considerable objection or even appear absurd. It begs the question of not only whether one's voluntary desire is necessary or sufficient for supporting hastened death, but also whether possession of rationality and/or self-consciousness is necessary for inclusion in the moral community. Without acknowledging that even those who have lost consciousness still possess dignity that is inviolable, Kant runs the risk of making these individuals disposable and denying any inherent value of caring for such humans.

Perhaps the relational aspect of personhood that is emphasized in Confucian philosophy can help us here.<sup>29</sup> While Kant and many Western philosophers emphasize the autonomy of a unitary moral subject as a central component of personhood, Confucians posit that human beings are by nature social, interdependent, and related to each other from the

---

<sup>29</sup> It is worth noting that many feminists also emphasize the relational nature of agency and personhood.

moment of birth.<sup>30</sup> Chinese Confucianism posits the family as the fundamental unit in society and emphasizes the importance of intergenerational and other interpersonal relationships as integral to one's personhood and identity. Individuals are not separate entities; they are socially situated and part of an interdependent network. Stressing the relatedness of personhood allows us to recognize that dependents whose reasoning capacities might have become severely limited by dementia and other compromising conditions are still contributing to the ongoing nature of human relationships even as they may no longer be able to recognize their loved ones.<sup>31</sup> What contributes to personhood here is not the patient's rational capacities, but his/her relationship with others. Such a conception enables us to recognize the importance of protecting and enhancing the central human bonds even as modern medicine cannot maintain or restore one's reasoning or self-consciousness. Discussions about assisted deaths that focus not on the patient's foregone autonomy, but on his or her relatedness can help us to explain why abandonment of such individuals may still violate their personhood.

---

30 E. Yu and R. Fan, "A Confucian View of Personhood and Bioethics", *Journal of Bioethical Inquiry* 4 (2007), 171–9; esp. 173.

31 J. Tao, "Confucian and Western Notions of Human Need and Agency: Health Care and Biomedical Ethics in the Twenty-First Century", in R. Qui (ed.), *Bioethics: Asian Perspectives* (Boston: Kluwer Academic Publishers, 2004), 13–28, esp. 24.



## 25. Human Dignity and the Innate Right to Freedom in National and International Law

*Bernhard Jakl*

### 1. Is There a Specific Legal Meaning of Human Dignity?

The concept of human dignity plays a crucial role in human rights discourse. Human dignity is a commonplace in national (e.g., in constitutional texts) and international law (e.g., in the UN Charter and several declarations and conventions). It is used frequently in court decisions all over the world, such as of the Supreme Court in the USA and the Federal Constitutional Courts in Germany and South Africa. The concept of human dignity enables different cultures with different conceptions of the state, differing views on the basis of human rights, and differing moral viewpoints to put aside these normative differences and agree on a basic minimum content of the meaning of human dignity: each human being possesses an intrinsic worth that should be respected.<sup>1</sup>

But beyond the cross-cultural political and juridical acceptance of this minimum content (or common core of human dignity) a more sound and coherent meaning of dignity is still missing. For example, it is quite unclear what the “intrinsic worth” of a human being stands for. If the diagnosis that human dignity is still a vague concept is correct, the question arises how human dignity can be interpreted as a *legal* term.

The purpose of this essay is to attempt a further exploration of the relation between the specific legal meaning of human dignity and Kant’s practical philosophy. First, I briefly present the German Constitutional Court’s interpretation of human dignity and one of its recent problems,

---

1 See e.g., the International Covenant on Civil and Political Rights, adopted and opened for signature, ratification and accession by General Assembly resolution 2200 A (XXI) of 16 December 1966 (entry into force 23 March 1976): “Recognizing that these rights derive from the inherent dignity of the human person ...”. For further examples see Christopher McCrudden, “Human dignity and judicial interpretation of human rights”, *European Journal of International Law* 19.4 (2008), 655–724; esp. 722–3.

the discussion whether human dignity is an absolute right. This discussion provides a prism for the problems one faces with interpreting human dignity as a right. Second, I analyze the relation between dignity and the innate right to freedom within Kant's practical philosophy. I will argue that, in order to provide crucial elements for a more specific meaning of human dignity in legal contexts, a Kantian interpretation must refer to the innate right to freedom of the Metaphysical First Principles of the Doctrine of Right rather than to the dignity of the *Groundwork of the Metaphysics of Morals*. Third, I conclude with a brief look at the effects this more specific meaning of human dignity has on law.

## 2. The German Interpretation: Human Dignity as Absolute Subjective Right

For at least two reasons the German discussion about the interpretation of human dignity seems to be an appropriate prism, if one wants to interpret human dignity as a right. First, the German constitutional court's jurisdiction adopts dignity as a *juridical* criterion for decision making. Second, the German legal discussion concerning the meaning of dignity refers to Kant. Kant's theory is seen worldwide as the basis of an enlightened, autonomy-based version of human dignity and therefore functions as a global frame of reference for discussing the meaning of human dignity.<sup>2</sup> Because of these "laboratory conditions" one can expect to identify elements for a more specific and coherent meaning of human dignity in legal contexts by confronting the German discussion on dignity with Kant's Critical philosophy of right.

Human dignity is the fundamental principle of the German constitution. Article 1, paragraph 1, says: "Human dignity shall be inviolable. To respect and protect it shall be the duty of all state authority." The recent trend in German discussion on dignity reveals a gap between jurisdiction and jurisprudence. The jurisdiction emphasizes the fundamental and absolute character of dignity.<sup>3</sup> If there is an infringement of dignity, no justification can be given. According to the German constitutional court, human dignity is violated, if someone is treated as a

---

2 McCrudden, 659, 724.

3 E.g., BVerfG, 1 BvR 357/05 of 15.2.2006: the Aviation Security Act (*Luftverkehrsgesetz*). (BVerfG refers to decisions of the German Federal Constitutional Court, followed by the file reference and date.)

mere object.<sup>4</sup> This so-called “Object-Formula” derives from one of Kant’s formulas of the categorical imperative, the so called “End-in-itself-formula”: “Act in such a way that you treat humanity, whether in your own person or in the person of any other, always at the same time as an end and never merely as a means to an end.”<sup>5</sup>

The Object-Formula is the criterion the German constitutional court based several decisions on. For example, §14, clause 3, of the “Aviation-Security-Act” (*Luftsicherheitsgesetz*) would have allowed the German federal army (*Bundeswehr*) to shoot down hijacked airplanes, if there was an expectation of the planes being used as weapons by terrorists. The law was declared unconstitutional on the grounds of human dignity: killing a small number of innocent people to save a large number cannot be legalized since it treats the innocent people (the smaller number) as mere objects for ensuring the safety of others (the larger number on the ground).<sup>6</sup>

In contrast to the jurisdiction, many academics criticize the idea of the Object-Formula as well as its consequence, the absolute character of dignity, as formal. Therefore the Object-formula is widely seen as not helpful for substantial legal decision-making. Most academics argue that dignity should be used as a relative, non-absolute right that can collide with other rights, such as the dignity, life, or property of others.<sup>7</sup> According to this academic opinion, it is possible to justify an infringement of dignity by balancing it with other rights. Consequently, dignity would lose its prominent and fundamental position based on its absolute character.

A summary of the German situation is as follows. On the one hand, the Kantian-inspired interpretation of the jurisdiction stresses the absolute character of human dignity, while dignity remains a vague concept, only defined by special cases concerning an infringement of dignity. On

---

4 Matthias Herdegen, “Kommentar zu Artikel 1 Abs. 1 Grundgesetz” (“Commentary on Article 1 of the German Constitution”), in M. Herdegen et al. (eds.), *Kommentar zum Grundgesetz (Commentary on the German Constitution)* (München: Maunz/Dürig, 2004); see Art. 1 Abs. 1 Rn. 28.

5 4:429; cf. Herdegen, Art. 1 Abs. 1 Rn. 28, 34.

6 BVerfG, 1 BvR 357/05 of 15.2.2006.

7 Herdegen, Art.1 Abs.1 Rn. 45; Horst Dreier (ed.), *Grundgesetz-Kommentar (German Constitution. A Commentary)*, second edition (Tübingen: Mohr Siebeck, 2007), Art.1 Abs.1 Rn. 133; Winfried Brugger, “Darf der Staat ausnahmsweise foltern?” (“Are there any exceptions so the state may torture?”), *Der Staat* 35 (1996), 67–97, esp. 67.

the other hand, the critics stress that a positive concept of dignity (going beyond the Object-Formula) is still missing, while denying the absolute character of dignity. This shows that a juridical use of the concept “human dignity” seems to be paradoxical: we can claim the absolute character of dignity, but we cannot identify its positive general meaning beyond an individual case. Or we can outline dignity as a standard right, like property, but as a result we have to relinquish its absolute character.

This situation gives rise to the question: how should we handle this paradoxical gap? One strategy is to ask whether one could combine a positive meaning of dignity with its absolute character. Since the problem originates from references to Kant’s categorical imperative, it seems appropriate to have a closer look at Kant’s ideas.

### 3. A “More Kantian Approach”: Human Dignity and Innate Right to Freedom

A Kantian approach to the meaning of “human dignity” has to take into account at least two different sources for Kant’s concept of personhood. First, the foundational level of the *Groundwork* and second the applicational level of the *Metaphysics of Morals*. After arguing in his first *Critique* that there is space for practical reasoning due to the limits of theoretical reasoning, Kant’s *Groundwork* is the starting point (and foundational level) for the positive version of his practical philosophy. The arguments presented in the *Groundwork* deal with the question: what ought I to do? After identifying the categorical imperative—freedom’s law for the area of practical reasoning—Kant offers different formulas to indicate how to apply the categorical imperative (4:421).

In one way or another all the formulas stress the Kantian point that every human being has the faculty as well as the duty to choose his or her ends freely. Kant’s theory of structured, lawful freedom stands as the point of reference of extensive debates that cannot be addressed here. Nevertheless, one can assume a kind of Kantian standard procedure present in all the discussions: first, sum up my own individual actions under more general practical hypotheses, the so-called maxims; second, submit the maxims to a generalization test by asking whether my maxims entail structures incoherent with their own premises.<sup>8</sup>

---

8 Cf. Andrea Esser, *Eine Ethik für Endliche. Kants Tugendlehre in der Gegenwart (Ethics under the Conditions of Finitude. Kant’s Doctrine of Virtue in the Present*

In this foundational context of the *Groundwork*, dignity is not a concept needed to argue for the categorical imperative. Dignity is a traditional concept used by Kant to exemplify his new idea that pure reason can be practical: “whatever ... is above all value, and therefore admits of no equivalent, has a dignity” (6:434). Dignity is a very broad cultural concept and Kant transforms its meaning within his practical philosophy. For Kant, dignity is no longer a concept of aristocratic exclusion but a concept of republican inclusion, because every being of human form can be seen as autonomous.<sup>9</sup> The meaning of dignity is now derived from autonomy instead of the other way around. Only because one can identify the moral law can one justify the idea of dignity as an absolute inner worth of every human being (4:436). But cultural concepts, even with a new meaning, are not juridical concepts. Furthermore, the idea to respect every human being for the reason that it is part of a structure whereby pure reason can in itself be practical seems indeed very formal.

Even if one admits that the Kantian standard procedure for applying the categorical imperative might be helpful as a guideline for individual decision-making and that dignity is an attribute of every human being, some important questions for lawyers remain unanswered, such as: What is the specific legal meaning of dignity? To whom should “dignity” be ascribed? Only to self-determined adults? What about children, elderly persons suffering from dementia, and disabled persons? Are all of them addressees of the categorical imperative in the same way?

The idea of autonomy, as an ideal and insofar necessarily formal standard expressed in the categorical imperative, differs from its realization under the conditions of finitude. Kant’s *Metaphysics of Morals* can be interpreted as the missing link between the idea of autonomy and its application in the real world, a world with different forms of external constraints for an autonomous life.<sup>10</sup> Within the practical field of norms, Kant draws a distinction between ethics and law. While the Doctrine

---

*Day*), *Spekulation und Erfassung II*, 53 (Stuttgart-Bad Canstatt: Verlag Frommann Holzboog, 2004), 281–92.

9 Cf. Oliver Sensen, “Kants Begriff der Menschenwürde” (“Kant’s concept of human dignity”), in F. J. Bormann und Chr. Schröder (ed.), *Abwägende Vernunft* (Berlin u. a.: de Gruyter, 2004), 220–36; esp. 231.

10 Cf. Esser, 252, and Günter Zöller, “Idee und Notwendigkeit einer Metaphysik der Sitten” (“Idea and Necessity of a Metaphysics of Morals”), in Andreas Trampota, Oliver Sensen, and Jens Timmermann (eds.), *Kants “Tugendlehre”* (Berlin/New York: de Gruyter, 2011).

of Virtue provides further information on the question of how an individual can live according to the moral law, the Doctrine of Right explores the question of how we can identify binding norms for a society of free individuals.

For applying the highest universal principles of the *Groundwork* to the objects of experience, the particular nature of the human being, Kant's project in the Doctrine of Right is to "supply the immutable principles for any giving of positive law" (6:229). For answering the legal question Kant transforms the abstract universal law of the categorical imperative to the more specific "universal principle of right": "Any action is right if it can coexist with everyone's freedom in accordance with a universal law" (6:230). According to Kant, law's empire is not defined by a catalogue of norms, rules of recognition, or substantial principles of humanity. From the standpoint of the universal principle of right, law must be seen as a never-ending coordination process of equal subjective rights of freedom.<sup>11</sup> This universal principle of right can be interpreted as a generic rule for the meaning of law as a cultural product. One can distinguish law from other normative fields, such as ethics. The object of the law is the coordination of different actions (or choices, *Willkür*). Action means to be "by means of one's representations (plans) the cause of the realization (object) of these representations (plans)".<sup>12</sup>

The focus on "action" excludes two other aspects from the legal standpoint of the doctrine of right: (a) whether pure reason can be practical by itself, as the categorical imperative (= *Wille*) suggests, is not a problem of legal reasoning (6:213); and (b) what someone wishes to do is not a problem concerning legal reasoning, until a person starts realizing his or her ideas through an action-plan. Such ideas (that someone wants to realize) then have to be called "actions" (= *Willkür*) (6:213).

Within the doctrine of right's juridical frame "there is only one innate right": "Freedom (independence from being constrained by another's action /choice), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right be-

11 Cf. Bernhard Jakl, *Recht aus Freiheit. Die Gegenüberstellung der rechtstheoretischen Ansätze der Wertungsjurisprudenz und des Liberalismus mit der kritischen Rechtsphilosophie Kants (Right from Freedom. A Confrontation of the Positions of Canaris and Dworkin in Legal Theory with Kant's Critical Philosophy of Right)*, *Schriften zur Rechtstheorie* (Berlin: Duncker & Humblot, 2009), 126, 131–7.

12 My amended translation of "...durch seine Vorstellungen Ursache der Gegenstände dieser Vorstellungen sein" (6:211).

longing to every man by virtue of his humanity” (6:237). The innate right to freedom demands ascribing the faculty to act to every being of human form—that is, the ability to realize his/her representations, the agenda he/she has in mind, as far as the actions/agenda can be coordinated with the actions/agenda of others. As Fichte puts it, a year earlier than Kant: individual human beings must be seen as causes of objects in the observable world.<sup>13</sup> Therefore, the adjective “innate” does not refer to a natural right in the sense of a given natural position. In contrast, the “innate right to freedom” is a necessary condition for the realization of law as a construction of reason because it is the connection between the “universal principle of right” as general rule and its concretization in a single case within a specific field of practical knowledge.<sup>14</sup>

From a Kantian point of view, the transfer of dignity into a legal context confuses the ethical question, “what ought I to do?”, with the more specific legal question, “what kind of norms can be binding for a society of free individuals?”. Kant’s differentiation between ethical rules for individual decision making and legal norms addressing the interaction between individual rights, as well as the different levels of foundation and application in Kant’s practical philosophy, constitute two strong systematic reasons for referring to the innate right to freedom as the universal right, if one wants to interpret human dignity in legal texts from a Kantian point of view. Human dignity as a legal concept thus turns out to be every human being’s subjective right: to be seen as someone who has the faculty to act.

---

13 Johann Gottlieb Fichte, *Grundlage des Naturrechts, nach den Prinzipien der Wissenschaftslehre* (*Foundations of Natural Right, according to the Principles of the Wissenschaftslehre*), in J. G. Fichte, *sämtliche/nachgelassene Werke*, ed. I. H. Fichte, Band III, 1–386 (1796), 113.

14 Cf. Jürgen Habermas, *Die Einbeziehung des Anderen. Studien zur politischen Theorie* (*The Inclusion of the Other. Studies in Political Theory*) (Frankfurt am Main: Suhrkamp, 1996), 225.

#### 4. Human Dignity as Innate Right to Freedom in the Law

Interpreting “human dignity” as innate right to freedom might not solve all the paradoxical problems related to human dignity within a juridical context. But it helps overcome the existing problems insofar as they originate from the missing, specifically *legal* meaning of dignity.

On the one hand, the innate right to freedom demands understanding human dignity in legal contexts as an absolute right that cannot be balanced with other rights. It is the normative core of the law, because it serves as a criterion for every giving of positive laws. Some things/norms cannot be implemented into law’s empire at all, because they are incompatible with the idea of an individual’s faculty to act, or with understanding an individual being of human form as an agent. Within law’s empire, one does not even have to discuss their justification, while from non-legal perspectives such a discussion is possible.

On the other hand, one can outline the universally accepted idea of human dignity a bit further. By referring to the faculty to act (i. e., the innate right to freedom) there is no longer any need to introduce non-legal values or a divine picture of personhood for a *legal* meaning of human dignity. The only function the innate right to freedom asks to be taken into account is the faculty to act. From a Kantian perspective, a specifically legal use of “human dignity” means to ask whether, for the persons concerned with a juridical norm or juridical action, a further realization of representations is still possible or not. If a norm denies the norm-addressees’ faculty to act (that is to realize representations), it infringes the innate right to freedom.

Because law does not only involve sharpening concepts but much more making decisions, I will briefly and finally return to the present German discussion. Here one realizes that critics of the jurisdiction interpreting dignity through the Object-Formula have a good point, but for a different reason. The problem is not the absolute character of human dignity. The problem is that using the Object-Formula to interpret human dignity not only results in disregarding the different levels of the Kantian theory (the *Groundwork* and the Doctrine of Right), but also confuses different fields of normative questions. Human dignity in legal contexts fails to deal with problems such as: What ought I to do? What are my maxims? What is in the best interest of a person or of a group? To this extent, only the implicit idea of the German Constitutional



Court's decision on the Aviation Security Act is relevant to the Kantian (and so also deontological) point: every individual has the absolute right to be seen as an agent. This right cannot be taken away legally in the name of a collective goal, nor can a person be downgraded to a non-agent for legal reasons. No legal norm taking away innocent citizens' capacity to have representations, as well as their capacity to realize their representations (or agenda), can be regarded as consistent with the general principle of right or the innate right to freedom. A positive law such as the Aviation Security Act would undermine the positive structure of already existing rights and destroy the normative idea of subjective rights within an existing juridical normative order. A situation such as the "trolley-scenario" (corresponding to the case of the Aviation Security Act) can exemplify the limits of legal reasoning from another (e. g., economic) perspective on a law regime, but cannot justify legal norms. The only two criteria to evaluate and justify a positive law, from the point of view of Kantian legal theory, are the general principle of right and the absolute subjective right to participate in this structure, the innate right to freedom.

## 26. “Irgend ein Vertrauen ... muss ... übrig bleiben”: The Idea of Trust in Kant’s Moral and Political Philosophy

*Peter Schröder*

This essay will explore the notion of trust and confidence in Kant’s writings. Our autonomy and thus also the organization of human society both depend on the possibility of mutual trust. This is a theme which seems less explored in Kant scholarship, but as I will argue figures prominently throughout Kant’s moral and political philosophy. His occasional criticism of Machiavelli (morality versus the doctrine of reason of state) and the Jesuits (“Jesuitencasuistik”) alike is just one rather obvious indicator that Kant fully realized the importance of trust. But this notion goes well beyond the classical idea of *pacta sunt servanda*. As a concept the idea of mutual trust holds an important and problematic place in Kant’s understanding of individuals and their relation to the state. Trust and its conceptual significance are paramount for Kant in forging a coherent theory of international relations. In discussing the different use and status of trust in both the relation between individual and state and international relations, we will be able fully to grasp the range and significance of the concept of trust in Kant’s moral and political philosophy.

Human nature is perceived by Kant as conflictual,<sup>1</sup> following Hobbes’ concept of a right to everything (*ius in omnia*) in the state of

---

1 See *The Metaphysics of Morals* (6:311) and *Toward Perpetual Peace* (8:341–86, esp. 355 and 376). Hobbes’ argument regarding the *ius in omnia* and the ensuing aporie of rights works on the same structural premise, but he does not distinguish with the same clarity as Kant and freely adds empirical arguments to foster his theory. See D. Hüning, *Freiheit und Herrschaft in der Rechtslehre des Thomas Hobbes* (Berlin: Dunker & Humblot, 1998). For a wider discussion see notably J. G. Fichte, “Machiavelli als Schriftsteller”, in *Werke*, vol. XI, ed. I. H. Fichte (Berlin: Walter de Gruyter, 1971), 420 (my translation): “The main principle of Machiavellian politics, and we may add without reservation, also of our own, and, in our opinion, indeed of any theory of the state, which understands itself, is contained in the following words of Machiavelli: ‘that in constituting and legislating for a commonwealth it must needs be taken for granted that all

nature. This is the first premise to take on board to appreciate fully the *problematique* that must be addressed in any discussion of trust. Thus coming from this same fundamental assumption as Hobbes, Kant pursues to a certain extent—and despite his negative rhetoric (8:273–313, esp. 289)—an argument that remains very similar to Hobbes'. Similar to the distinction Hobbes draws between the *foro interno* and the *foro externo*,<sup>2</sup> Kant distinguishes between duties of right—as is well known, he declared these in his *Metaphysics of Morals* to be perfect duties—and duties of virtue or imperfect duties, belonging to the sphere of ethics (6:239 f). To the perfect duties of right corresponds the legal power of the lawgiver, so that these duties can be enforced in the case of non-compliance. Kant asserts (6:219 f) that

ethics commands that I still fulfill a contract I have entered into, even though the other party could not coerce me to do so; but it takes the law (*pacta sunt servanda*) and the duty corresponding to it from the doctrine of right, as already given there. Accordingly the giving of the law that *promises agreed to must be kept* [my emphasis] lies not in ethics but in *Ius*. All that ethics teaches is that if the incentive which juridical lawgiving connects with that duty, namely external constraint, were absent, the idea of duty by itself would be sufficient as an incentive .... It is no duty of virtue to keep one's promises but a duty of right, to the performance of which one can be coerced.

This is not to imply that Kant maintained the position that keeping one's promise (i. e., being a trustworthy person) does not matter for individuals. What he clearly saw, however, was that the notion of trust or keeping one's promise has a different significance within civil society, where there is a lawgiver endowed with legitimate coercive power, than in a situation where precisely this framework is lacking, as is the

---

men are wicked and that they will always give vent to the malignity that is in their minds when opportunity offers'. [*Discorsi* I.3, 111 f.] It is not even necessary to discuss the question, whether this view of human nature does correspond to reality as assumed in this sentence, or not. In short, the state, as a constraining authority, does suppose humans to be thus and only this supposition justifies the existence of states." (There exists to my knowledge no English translation of this important writing by Fichte).

- 2 Cf. T. Hobbes, *On the Citizen*, ed. R. Tuck (Cambridge: Cambridge University Press, 1998), III.27, 53 f. Hobbes does not, of course, reach Kant's level of conceptual clarity and distinction between ethics and rights, but his idea of the obligatory character of natural laws in the state of nature clearly informed Kant's understanding. Cf. G. Geismann, "Kant als Vollender von Hobbes und Rousseau", *Der Staat* 21 (1982), 161–89.

case in the relation between sovereign states. In anticipation of the way Kant further develops his argument, especially regarding his theory of international relations, we can already see at this point that because there is no external lawgiver above the individual sovereign states, it is also inconceivable to Kant that there could exist perfect, and thus legally enforceable, rights or duties among nations. This argument is also clearly informed by Hobbes's premise of the paradoxical *ius in omnia* in the state of nature: "Nations, as states," Kant argues in *Toward Perpetual Peace* (8:354 f),

can be appraised as individuals, who in their natural condition (that is, in their independence from external law) already wrong one another by being near one another; and each of them, for the sake of its security, can and ought to require the others to enter with it into a constitution similar to a civil constitution, in which each can be assured of its right.

But, still following Hobbes,<sup>3</sup> Kant restricts this analogy (8:355), because

what holds in accordance with natural right for human beings in a lawless condition ... cannot hold for states in accordance with the right of nations (since, as states, they already have a rightful constitution internally and hence have outgrown the constraint of others to bring them under a

---

3 Hobbes had already restricted the analogy between the state of nature and states facing each other in *Leviathan*, where he wrote (T. Hobbes, *Leviathan*, ed. R. Tuck (Cambridge: Cambridge University Press, 1991), XIII.12, 78): "But though there had never been any time wherein particular men were in a condition of war one against another, yet in all times kings and persons of sovereign authority, because of their independency, are in continual jealousies, and in the state and posture of gladiators, having their weapons pointing, and their eyes fixed on one another; that is, their forts, garrisons, and guns upon the frontiers of their kingdoms, and continual spies upon their neighbours, which is a posture of war. But because they uphold thereby the industry of their subjects, there does not follow from it that misery which accompanies the liberty of particular men." Hobbes thus still maintains that interstate-relations resemble the state of nature of individuals, but the former is nevertheless a more stable condition since it allows for interior peace, security and industry. Cf. P. Schröder, "Natural Law, Sovereignty and International Law: A Comparative Perspective", in I. Hunter and D. Saunders (eds.), *Natural Law and Civil Sovereignty* (Houndmills: Palgrave MacMillan, 2002), 204–18. See also G. W. F. Hegel, *Elements of the Philosophy of Rights*, ed. A. W. Wood (Cambridge: Cambridge University Press, 1991), 368, §333: "The principle of international law ... should be observed. But since the sovereignty of states is the principle governing their mutual relations, they exist to that extent in a state of nature in relation to one another."

more extended law-governed constitution in accordance with their concepts of right).

But I do not intend here to pursue this well researched aspect of Kant's theory any further.<sup>4</sup> Instead I would like to emphasize again and further explore the importance and conceptual significance of trust, especially in Kant's theory of international relations—an aspect that seems generally ignored in those studies discussing Kant's theory of international relations.<sup>5</sup> That our behavior should be guided by the ethics developed in the *Metaphysics of Morals* seems clear (6:215 f). We might feel or hold trust toward others, and our own behavior should justify the trust others have in us. If our trust is proved unjustified by the behavior of the person whom we trusted, we will be disappointed. The other has thus not fulfilled his imperfect ethical duty. It seems to me that on a purely individual level Kant's ethical demand of the categorical imperative would demand us to be trustworthy and not to lie. But if we do lie and are thus clearly untrustworthy, it does not necessarily follow that we will be sanctioned beyond the individual or private sphere of disappointment or reproach.<sup>6</sup> Only if this breach of trust concerns at the same time a breach of a perfect duty of right by the other can we apply to the state authority with a claim to have been wronged unduly by the one

- 
- 4 The collections of essays in the following volumes are very pertinent on this aspect: R. Merkel and R. Wittmann (eds.), *„Zum ewigen Frieden“. Grundlagen, Aktualität und Aussichten einer Idee von Immanuel Kant* (Frankfurt/Main: Suhrkamp Verlag, 1996); and D. Hüning and B. Tuschling (eds.), *Recht, Staat und Völkerrecht bei Immanuel Kant* (Berlin: Duncker & Humblot, 1998).
- 5 For a more general approach regarding the importance of trust for a theory of international relations, see the illuminating article by P. Delholm, “Das Wagnis des Vertrauens”, in A. Hirsch and P. Delhom (eds.), *Denkwege des Friedens Aporien und Perspektiven* (Freiburg/München: Verlag Karl Alber, 2007), 334–61. I have treated the relevance of trust for inter-state relations in early modern political thought in much more detail in my “*Sine fide nulla pax*—Überlegungen zu Vertrauen und Krieg in den politischen Theorien von Machiavelli, Gentili und Grotius”, in M. Formisano and H. Böhme (eds.), *War in Words: Transformations of War from Antiquity to Clausewitz* (Berlin: de Gruyter, 2010), 37–60, and “Taming the Fox and the Lion—some aspects of the sixteenth century's debate on inter-state relations”, in O. Asbach/P. Schröder (eds.), *War, the State and International Law in Seventeenth Century Europe* (Farnham: Ashgate, 2010), 83–102.
- 6 In “On a supposed right to lie from philanthropy” Kant says (8:427): “I here prefer not to sharpen this principle to the point of saying: ‘Untruthfulness is a violation of a duty to oneself.’ For this belongs to ethics, but what is under discussion here is a duty of right.”

who did not keep his promise. On these grounds we can make a demand to the state for arbitration on this matter, with the expectation that judgment will address the wrong we received because of the experienced breach of our trust. On the other hand, the same construction of the argument (i. e., the issue of lying within the framework of legal civil society) allows Kant to insist (8:427) that “one who *tells a lie*, however well disposed he may be,<sup>7</sup> must be responsible for its consequences even before a civil law court.” Where the legal framework of civil society is in place, Kant can, therefore, argue (8:427) that “truthfulness is a duty that must be regarded as the basis of all duties to be grounded on contract, the law of which is made uncertain and useless if even the least exception to it is admitted.” On this basis (8:427): “to be *truthful* (honest) in all declarations is therefore a sacred command of reason prescribing unconditionally, one not to be restricted by any conveniences.” It is important to stress, though, that Kant's discussion of trust within civil society does not end here.

In his discussion of, or rather against, a right of resistance, Kant unconditionally prohibits any such right and clearly states (8:299 f; see also 6:320):

even if [the supreme legislative] power or its agent, the head of state, has gone so far as to violate the original contract and has thereby, according to the subject's concept, forfeited the right to be legislator inasmuch as he has empowered the government to proceed quite violently (tyrannically), a subject is still not permitted any resistance by way of counteracting force. The ground of this is that in an already existing civil constitution the people's judgement to determine how the constitution should be administered is no longer valid. For suppose that the people can so judge, and indeed contrary to the judgement of the actual head of state; who is to decide on which side the right is?

Kant was well aware that thinkers such as Achenwall, whom he cites (8:301), conceived of a right of resistance on the grounds that, since

---

7 Kant refuted the example that I might well lie if a murderer pursues my friend who is hiding in my house and asks me whether my friend is hiding in my place with the remark that I cannot know what might actually ensue from my lie, as for example, my friend might be just about to leave the house and might run into the murderer because I told him the lie that nobody was hiding here, and the murderer might thus leave at the same time as my friend, whereas if I had told the truth the murderer would presumably have searched the house and my friend would thus have had ample time to get away. As Kant adds (8:427), “if you have kept strictly to the truth, then public justice can hold nothing against you, whatever the unforeseen consequences might be.”

the head of state did not keep his part of the contract (i. e., did not keep his promise) and thus was not trustworthy any longer (6:273), there was no reason for the people to keep their part of the contract either. Kant's categorical denial of any right of resistance takes the notion of mutual trust into account. This seems to indicate that Kant was well aware of the importance of such a mutual relationship of trust between the people and their head of state. But he forestalls the kind of argumentation as seen in Achenwall by claiming exactly the opposite: on the grounds of the existing civil public constitution the sovereign has the legitimate right on his side. Only if the people had enshrined a right of resisting the head of state already in the constitution could they claim legitimacy for their resistance, but for Kant this "is an obvious contradiction" (8:303). Thus he uses the argument of mutual trust in connection with the idea of publicity to preclude any legitimacy of resistance, because (8:303):

no right within a state can be concealed, treacherously as it were, by a secret reservation, least of all the right that the people claims for itself as one belonging to the constitution; for all laws of the constitution must be thought as arising out of a public will. Thus if the constitution permitted insurrection, it would have to declare publicly the right to it and in what way use is to be made of it.

There can be no doubt, once a civil constitution is in place, the legitimate sovereign and arbiter, endowed with coercive power, cannot be overturned or even challenged by the use of power. The only way of addressing a wrong lies in public criticism, "the sole palladium of the people's rights" (8:304). But what are the implications for international relations where the sovereign states are precisely lacking such a legitimate arbiter, who is endowed with coercive power? Here the concept of trust in connection with the idea of publicity re-emerges in a very different light and receives a much more crucial conceptual significance in Kant's theory of international relations.

The situation between sovereign states does not allow for a concept of right, because (8:356 f):

the concept of the right of nations as that of a right to go to war is, strictly speaking, unintelligible (since it is supposed to be a right to determine what is right not by universally valid external laws limiting the freedom of each but by unilateral maxims through force) ... In accordance with reason there is only one way that states in relation with one another can leave the lawless condition, which involves nothing but war; it is that, like individual human beings, they give up their savage (lawless) freedom, accommodate them-

selves to public coercive laws, and so form a ... *state of nations* ... that would finally encompass all the nations of the earth. But, in accordance with their idea of the right of nations, they do not at all want this, thus rejecting *in hypothesi* what is correct *in thesi*; so (if all is not to be lost) in place of the positive idea of a *world republic* only the *negative* surrogate of a *league* that averts war ... can hold back the stream of hostile inclination that shies away from right, though with constant danger of its breaking out.

In this context we can now see the significance of Kant's conceptual use of trustworthiness. Kant clearly reaches the point in his argument on international relations where he cannot argue in the strict sense of his theory of right, since he rules out the solution of a sovereign lawgiver and arbiter himself. What he proposes is thus a second best solution that necessarily compromises his strict notion of right. Any alternative that avoids a solution in analogy to his theory of right necessarily has to be a surrogate. The merit of Kant's theory in having reached this point cannot be overestimated; indeed, contemporary discussion about the organization of international relations still reaches the same dead-end when it comes to the question of sovereignty regarding the relations of states.<sup>8</sup> Even Kant's alternative of progressively approaching a free federation of (republican) states remains a challenging task and is in many respects far from being achieved.<sup>9</sup> One way of engaging with this discussion would be to look at the arguments Kant provided in order to foster the second best alternative of a free federation. His notion of trust and the way he conceptually linked this idea with his notion of the public sphere are indeed a crucial prerequisite to safeguarding the

---

8 Geoffrey Vaughan has recently shown that modern political thought is found crucially lacking when it comes to addressing the whole question of sovereignty. See G. Vaughan, "The Decline of Sovereignty in the Liberal Tradition: The Case of John Rawls", in M. Peters and P. Schröder (eds.), *Souveränitätskonzeptionen. Beiträge zur Analyse politischer Ordnungsvorstellungen im 17. bis zum 20. Jahrhundert* (Berlin: Dunker & Humblot, 2000), 157–85. John Rawls argued against the use of sovereignty in his theory of international relations and claimed to follow Kant in doing so. See J. Rawls, *The Law of Peoples* (Cambridge, Ma.: Harvard University Press, 1999). Hans Kelsen already explicitly stressed for very similar reasons as Rawls that "the concept of sovereignty is to be radically abolished"; H. Kelsen, *Das Problem der Souveränität und die Theorie des Völkerrechts. Beitrag zu einer reinen Rechtslehre* [1928] (Wien: Franz Deuticke, 1960), 320.

9 See W. Kersting, "Weltfriedensordnung und globale Verteilungsgerechtigkeit. Kants Konzeption eines vollständigen Rechtsfriedens und die gegenwärtige politische Philosophie der internationalen Beziehungen", in Merkel and Wittmann, 172–212.



volatile concept of a federation that leaves ultimate sovereign rights with individual states. In conclusion I want to scrutinize the central role Kant attributed to the concept of trust regarding international relations.

The sixth and last Preliminary Article in *Perpetual Peace* states (8:346): “No state at war with another shall allow itself such acts of hostility as would have to make mutual trust impossible during a future peace.” This is easily overlooked and underestimated and I would like to suggest that trust or good faith plays a considerable and crucial role in Kant’s concept of how to work progressively toward perpetual peace. The concept of trust is not exhausted with the admittedly important claim (8:346) that “some trust in the enemy’s way of thinking must still remain even in the midst of war, since otherwise no peace could be concluded and the hostilities would turn into a war of extermination.” The notion of trust and publicity are indeed the concepts Kant reintroduces as a “*transcendental formula of public right*” (8:381) even in relation to the “lawless” (8:357) sphere of sovereign states facing each other; this, as we have seen, had been the crucial point leading Kant to deny the existence of any notion of right between sovereign states (8:357). Kant claims the “*transcendental formula of public right*” for international relations is: “All actions relating to the rights of others are wrong if their maxim is incompatible with publicity” (8:381). This principle permits one only to determine what is not right toward others. But on this basis mutual trust is possible since there is a clear criterion that allows one to rule out any secret machinations. For Kant this is not just an ethical question but also an issue for the juridical sphere. What is still lacking, of course, is the possibility to seek arbitration by a higher power. In this respect Kant freely acknowledged that in the end states remained in the unlawful state of nature. What he clearly provided was an indication of how to make it less volatile. This remained for him “a task, that gradually solved, comes steadily closer to its goal” (8:386). For this gradual progress toward perpetual peace the possibility of mutual trust was one crucial condition. Kant safeguarded this condition by introducing the concept of publicity. As in his writing *An answer to the question: What is enlightenment?* the idea of progress depends on the condition of critical publicity, and this in turn creates the framework or atmosphere for mutual trust.

PART VI

Persons in Politico-Cultural  
Community



## 27. Autocracy: Kant on the Psycho–Politics of Self-Rule<sup>1</sup>

Günter Zöllner

“He was accustomed to calling his upper and lower faculties of the soul the Upper House and Lower House, and very often the former passed a bill that the latter rejected.”<sup>2</sup>

### 1. The Autonomy of the Will

In the *Republic* Plato famously draws an extended analogy between the set-up of the soul (ψυχή) and that of the city state (πόλις), relying on the latter to illuminate the former and thereby deriving key concepts and crucial distinctions of ethics from political philosophy. On the face of it, Kant’s practical philosophy might seem to offer a completely different and even opposed philosophical account of the relations that constitute the self. To begin with, Plato is not a major reference in Kant’s thinking about the constitution of the practical self (and neither is Aristotle). The ancient background for Kant’s ethics is Stoicism and hence an individualistic ethics rather than the integration of ethics into a larger, essentially political conception of worthy human life to be found in Plato (and in Aristotle). Kant’s reception of Plato’s *Republic* seems limited to the latter serving as a paradigm for the regulative function of a “concept of reason” (*Vernunftbegriff*) or “idea” (*Idee*)—that is, an infinitely removed but orienting as well as motivating “arch-image” (*Urbild*) that finite human endeavors are to strive after without ever reaching it (see

- 
- 1 A German version of this essay has been incorporated into my essay, “Autokratie. Die Psycho–Politik der Selbstherrschaft bei Platon und Kant”, in Hubertus Busche and Anton Schmitt (eds.), *Kant als Bezugspunkt philosophischen Denkens*, (Würzburg: Königshausen & Neumann 2010), 351–77.
  - 2 Georg Christoph Lichtenberg, *Schriften und Briefe*, Wolfgang Promies (ed.), 4 vols. and a volume of commentary (Munich: Hanser, 1967–74), vol.1, 65 (*Sudelbücher/ Waste Books*, B67).

B370–7/A316–20). With respect to practical ideas, such as “the Platonic Republic” (*die platonische Republik*), the striving is itself practical or a matter of human action. By contrast, speculative ideas, chiefly those of the soul, the world as a whole, and God, are the object of infinite striving in cognitive approximation through the acquisition of ever more extensive and intensive knowledge.<sup>3</sup>

Moreover, Kant’s political and legal philosophy seems entirely unrelated to the ideal or real accounts of life in the city state to be found in Plato (and Aristotle). The historical context of Kant’s political and legal philosophy is the sovereign territorial state of the modern era. Its theoretical background is the realist and idealist accounts of the nature of statehood in Machiavelli and Hobbes and in Locke and Rousseau and especially the modern tradition of praeter-positive, “natural law” (*Naturrecht*) or “law of reason” (*Vernunftrecht*).

Most importantly, at the center of Kant’s practical philosophy lies a conception of freedom as the capacity for absolutely spontaneous inner and outer action that is alien to the ancient world in general and to Plato’s (and Aristotle’s) outlook on the human being in particular. In ancient Greek and Roman philosophy human thinking and acting is considered an integral part of the world’s natural order and consists, to a large extent, in discovering and heeding the cosmic order. Even the philosopher’s vision of the Forms in Plato, the life of quasi-divine contemplation in Aristotle, or the self-sufficient life of the wise one in Stoicism does not depart from the natural order but represents supreme modes of conforming to it.

By contrast, Kant’s practical philosophy is grounded in a conception of freedom as anti-nature. The prerequisite for this radical reversal in the relation between the human being and nature is a changed understanding of nature and of the place of the human being in it. The scientific revolution of the early modern period had replaced the teleological cosmology of the ancients, chiefly preserved and passed down in Aristotle’s *Physics*, with the mechanistic image of a world made up of matter and governed by universal laws that lend themselves to mathematical repre-

---

3 On the possible transmission of Plato’s theory of forms to Kant through Moses Mendelssohn’s dialogue, *Phaedo, or the Immortality of the Soul*, see Klaus Reich, *Kant und die Ethik der Griechen* (Tübingen: Mohr, 1935). On the general significance of ancient philosophical theories for the development of Kant’s Critical philosophy, see Ulrike Santozki, *Die Bedeutung antiker Theorien für die Genese und Systematik von Kants Philosophie. Eine Analyse der drei Kritiken* (Berlin/New York: Walter de Gruyter, 2006).

sentation. Modern philosophers from Descartes through Leibniz to Kant contributed to the philosophical foundations of the new scientific worldview and sought to address its implications and consequences for a specifically modern understanding of human existence. In assigning the natural world, consisting of objects in space and time that stand under universal laws of nature, to the domain of “appearances” (*Erscheinungen*), Kant intentionally left vacant the realm of things as they are in themselves (*Dinge an sich, Dinge an sich selbst*). While entirely inaccessible to theoretical cognition (*Wissen*) and constitutively limited to the (scientific) knowledge of spatio-temporal appearances, the intelligible or noumenal world (*Noumena, mundus intelligibilis*) provided a conceptual space for the mere thought, as opposed to the determinate, intuitively warranted knowledge, of non-natural or free beings and their dynamic interactions in terms of substantiality, causality, and reciprocity. Kant went on to argue that, although the rational concept of absolutely unconditioned causation or “cosmological freedom” (*kosmologische Freiheit*) remained a merely regulative idea in the sphere of theoretical cognition, it assumed a reality of its own as the practically real idea, as “practical freedom” (*praktische Freiheit*), in the sphere of human moral agency (see A532–7/B560–5 and A801 f/B829 f).

A further feature of Kant’s radical reconceptualization of nature and freedom in the aftermath of the scientific revolution is the correlation of the newly established supra-natural freedom, that defines the human being in opposition to natural determination, with laws of its own, different from the laws of nature accounted for by modern natural philosophy and discovered by experimental physics. The strict character of the newly found or discovered “laws of freedom” (*Gesetze der Freiheit*) resembles the universal and necessary validity of the laws of nature. But on Kant’s account, the laws governing freedom are unlike the laws governing nature in that the latter are followed unfailingly and automatically, while the former are followed only contingently. To be sure, that the laws of freedom are followed only contingently and due to further conditions on the part of the beings following them (human beings), for Kant, is compatible with the unconditional status of these very laws as necessarily valid norms of human conduct. Laws of freedom have validity even if they happen not to be valued. In fact, the modal difference between the laws of freedom being valid strictly necessarily and their being followed only contingently is indicative of the complex constitution of the practical self in Kant’s mature moral philosophy.

Like Plato before him, Kant resorts to political conceptuality to cast and characterize the specifics of the structure of the practical self. While Plato's account of the inner human being takes recourse to the city state as the human being writ large (μακροάνθρωπος), Kant's theory of practical subjectivity is conceived under the formative influence of modern political thought. In particular, Kant casts the constitution of the practical self and its relation to the laws of freedom in terms of the political powers that constitute the modern state and of the origination and efficacy of its laws. On Kant's account, the practical self stands to the laws of freedom in a twofold relation that can be expressed by the juridico-political distinction between the giving-out and the carrying-out of laws (i. e., of legislation and execution). This distinction goes back to Montesquieu's influential analysis of the origin and function of political laws in relation to the powers (*pouvoirs*) that constitute the state in *The Spirit of the Laws*.<sup>4</sup> Montesquieu's historically informed reflections on the formative interrelations between law and politics subsequently were reshaped into a normative account of the contractual origin and quasi-democratic exercise of state power in Rousseau's *On the Social Contract*, where legislative and executive authority (*puissance législative*, *puissance exécutive*) were distinguished from each other and related to each other as will (*volonté*) and power (*pouvoir*).<sup>5</sup>

Kant, drawing on Rousseau, distinguishes in the individual self—more precisely, in the practical individual self—the sovereign that gives the law and the subject that is to follow it. In Rousseau's theory of the sovereignty of the people the same political body that has the power to issue laws (viz., the people) is also the political body the laws are addressed to. Analogously, in Kant the practical self is both the sovereign lawgiver and the subject of obedience with respect to the laws of freedom. Kant also follows Rousseau's political philosophy in identifying the capacity involved in the act of legislation. In Rousseau this is the will, more precisely the legislative will or the “general will”

---

4 See Montesquieu, *De l'esprit des lois*, ed. Robert Derathé, 2 vols. (Paris: Garnier, 1973), vol.1, 168 (Book XI, ch. VI). Montesquieu here characterizes jurisdiction as a kind of executive power: the executive power regarding matters that depend on civil law (*droit civil*), as opposed to the executive power regarding matters that depend on public law (*droit des gens*).

5 See Jean-Jacques Rousseau, *Du contrat social*, ed. Pierre Burgelin (Paris: Garnier, 1966), 97 (Book III, ch. 1).

(*volonté générale*).<sup>6</sup> In Kant the practical laws are issued by the will, more precisely by the “pure will” (*reiner Wille*) (4:390; 5:30).

Kant also takes over the political term for the sovereignty of the lawgiving will, “autonomy” (*Autonomie*), from political discourse into moral philosophy (4:440), retaining the apersonal, institutional understanding of the term. In Kant’s practical philosophy autonomy in the strict, technical sense of independent legislative authority does not pertain to the practical self as such but only to the will as the faculty of lawgiving, in contradistinction from those functions, capacities, or faculties that enable the self-given law to be—or not to be—followed. Accordingly, the will in its legislative capacity as pure will—as morally pure and purely moral will—strictly and properly speaking is not free. Nor is it unfree (6:226). Rather, it is to be considered as the source of the laws of freedom or of those laws that govern the free exercise of the practical capacities and capabilities of the self.

Accordingly, Kant distinguishes between the legislative “will” (*Wille*) and the executive “faculty of choice” (*Willkür*): only the latter can be considered free in beings like us (human beings), who are capable of rational, reason-based conduct on the basis of deliberative choice (6:213 f, 226 f). By contrast, non-rational finite beings or brutes may operate on the basis of choice between competing instincts or impulses. But the choice is internally necessitated, by laws of nature, and does not involve a “faculty of free choice” (*freie Willkür*) (6:213 f; see also A801 f/B829 f). Moreover, beings like us, capable of free choice, are not only able to choose their conduct freely on the basis of practical reasoning, or reasoning about reasons for action in general. They are also able to act specifically on the basis of laws of freedom alone, unaided by further considerations and reasons. Kant terms the capability of finite rational beings to act not only on the basis of reasons of all kinds but specifically and exclusively on the basis of reason alone “pure practical reason” (*reine praktische Vernunft*) (4:389; see also 5:3, 30). Moreover, on Kant’s quasi-political account of the matter, pure practical reason is both the legislative authority and the executive authority in establishing the laws of freedom.

The basic type underlying the manifold laws of freedom countenanced by Kant is that of the lawfulness of pure practical reason as such, consisting in the sheer form of universality. Laws of freedom realize freedom by restricting the latter to the formal condition of univer-

---

6 See Rousseau, 54 (Book I, ch. 7), 63 f (Book II, ch. 1).



sality. In subordinating the free exercise of willing to an order of conduct, laws of freedom substitute “wild” freedom with “ordered freedom” (*geordnete Freiheit*).<sup>7</sup> Kant employs the traditional term “morals” (*Sitten*), corresponding to the Latin *mores* (customs), to designate the ordered character of human conduct. But the plural form taken over from tradition should not be taken to indicate an ultimate plurality of orders regulating human conduct depending on time and place. Rather, Kant unites the plural rules of conduct under the conception of a single, original “law of morals” (*Sittengesetz*) or “moral law” (*moralisches Gesetz*) (4:38, 41).

The plurality of laws of freedom recognized by Kant, then, is not a manifestation of a pluralistic conception of competing morals but results from the further specification of the single “law of morals” that posits the form of rational universality as such into plural, contextually specific laws of conduct. Moreover, the content and hence the plurality of moral laws do not originate in pure practical reason *per se*; the latter is essentially limited to the presentation and imposition of mere rational form (universality) to possible or actual principles of conduct that may originate outside and independent of reason proper. Kant terms such principles of conduct that govern the conduct of individuals prior to the formal supreme law of morals “maxims”, resorting to the traditional term for maximally comprehensive rules of conduct (4:27). Strictly speaking, then, the autonomy of the will, conceived by Kant along the political lines of legislation in the state, is limited to the meta-principle that conduct according to maxims has to be susceptible to take on the form of universal legislation. Moral conduct, however individually determined by given maxims, is to satisfy the minimal formal condition that the given maxims can be the object of universal legislation.

Having cast the origin and status of the moral law in terms of the political institution of state legislation, Kant goes on to characterize the status of the moral law with respect to the human being by means of another feature of modern political theory and practice: the bicameral system of legislation, with different bodies cooperating or competing to draft and pass laws. Kant views the practical self as constituted by two basically different modes for determining its conduct. The generic “faculty of desire” (*Begehrungsvermögen*), whereby “representations” (*Vorstellungen*) operate as grounds for action, is divided into a

---

7 19:276;R7202; see also Wolfgang Kersting, *Wohlgeordnete Freiheit. Immanuel Kants Rechts- und Staatsphilosophie* (Frankfurt/M.: Klostermann, 1993).

“lower” and “higher” faculty of desire (*unteres Begehrungsvermögen*, *oberes Begehrungsvermögen*) (4:9n; 6:211–4; 5:178). While the higher faculty of desire is susceptible of purely rational grounds for action, the lower faculty of desire has its grounds for action provided by non-rational, sensory factors.

In view of the fact that in rational practical beings like us (humans) the sensory grounds for action do not by themselves constitute a sufficient reason for action but are subject to approval—or disapproval—by the “faculty of free choice” (*freie Willkür*), the determining grounds of the lower faculty of desire are to be considered as “inclinations” (*Neigungen*). Their influence inclines the practical self without necessitating it. The “faculty of choice” (*Willkür*) can be called “free” insofar as it is independent from internal necessitation by prevailing desires (A801 f/ B829 f). Yet while such desires do not necessitate the faculty of choice, they still influence and “affect” (*affizieren*) it and thereby may lead it to decline the requirements of the moral law.

Just as the faculty of free choice is not necessarily determined by desires, it does not necessarily follow the moral law either. In order to be able to prevail over competing inclinations as the sufficient determining ground of the will, the moral law as issued by legislative will has to present itself to the finite rational being as commanding adherence, and as unconditionally commanding it at that. The moral law appears under the quasi-political guise of an executive order instructing the practical self to follow it under all circumstances. The possible universality of maxims implied by the moral law thus takes on the outward form of the categorical imperative to act on those and only on those maxims that are susceptible to universal legislation (4:414–9; 5:19–21).

Kant explains the intensional difference between the moral law and the categorical imperative in extensional terms: the moral law holds for all finite rational practical beings, including being like us (humans) but also morally perfect beings, if there are any, while the categorical imperative pertains only to those finite rational practical beings that are subject to inclinations contrary to the requirements of the moral law and that hence are inclined, although not necessitated, not to follow it. Given the unconditional commanding character that the moral law takes on under conditions of finite, sensorily affected practical rationality, the categorical imperative functions not only as the principle for the cognition of how human beings are to conduct themselves (*principium dijudicationis*), but also as the principle for the recognition, or acknowledgment, of such conduct; in the latter role, it serves to motivate the execution

of the moral law (*principium executionis*).<sup>8</sup> Modally speaking, the “necessity” (*Notwendigkeit*) of the moral law takes on the further character of a “necessitation” (*Nötigung*) (5:32) that represents what is practically necessary but may not occur, due to countervailing and prevailing inclinations as required and commanded by the moral law or as “duty” (*Pflicht*) (4:397, 400).

## 2. The Autocracy of the Self

In Plato’s *Republic* what later came to be differentiated terminologically as ethics (ἠθικῆ) and political science (πολιτικῆ) found a unitary treatment attesting to the thoroughly political character of ethics in classical ancient thought, prior to the depolitization and privatization of ethics in Hellenistic Greece as well as Republican and Imperial Rome. Even Aristotle, who introduced the disciplinary distinction of practical philosophy into ethics, politics, and economics, treats ethics as an integral part of a comprehensively conceived study of political matters and has the *Nicomachean Ethics* lead over into the wider field of things political.<sup>9</sup> The unitary conception of practical philosophy in Plato (and Aristotle) finds a late modified continuation in Kant, for whom “practical philosophy” (*praktische Philosophie*) coincides with “moral philosophy” (*Moralphilosophie*): he excludes economics and other fields of technical knowledge that put theoretical cognition to practical use, because they appeal to laws of nature without relying on genuinely practical principles that involve laws of freedom (5:171; 6:217 f).

In a manner reminiscent of the ancient divide of practical philosophy into ethics and politics Kant distinguishes “right” (*Recht*) and “ethics” (*Ethik*) as the two parts of practical philosophy. Yet unlike Plato and Aristotle, who had limited *a priori* knowledge to the natural and supernatural objects of theoretical philosophy (τὰ φυσικά, τὰ μετὰ τὰ φυσικά), Kant maintains the susceptibility and even the need of practical philosophy to specify a body of non-empirical practical knowledge (6:214–8). The latter takes the form of *a priori* principles of law and ethics or of “the metaphysics of morals in two parts” (*Die Metaphysik der Sitten in zwei Teilen*), consisting of “Metaphysical First Principles of

8 See Immanuel Kant, *Lectures on Ethics*, tr. Louis Infield (Indianapolis/Cambridge: Hackett, 1963), 36; Immanuel Kant, *Vorlesung zur Moralphilosophie*, ed. Werner Stark (Berlin/New York: Walter de Gruyter, 2004), 55 f.

9 *Nik. Eth.* 1180a–1181b.

the Doctrine of Right” (*Metaphysische Anfangsgründe der Rechtslehre*) and “Metaphysical First Principles of the Doctrine of Virtue” (*Metaphysische Anfangsgründe der Tugendlehre*) (6:205 f, 375–7).

The unity of practical philosophy, and specifically that of the two-part metaphysics of morals, is assured by the shared function of the categorical imperative as the supreme principle of law and ethics. A further generic feature of practical philosophy, regardless of the difference between law and ethics, is its concern with a “legislation” (*Gesetzgebung*) that involves laws of freedom or laws that make possible the realization of freedom (6:218). The specific difference between juridical and ethical legislation comes in through the parallel distinctions between “inner” and “outer actions” (*innere, äußere Handlungen*) and between legislation that makes only the action itself a duty and legislation that, in addition to making the action a duty, also makes an action’s being a duty the motivating force (“incentive”, *Triebfeder*) for the action. Refraining from motivational prescriptions, the former legislation is juridical and involves all those and only those obligations (“duties”; *Pflichten*) that can be legislated externally, through one or more persons or an institution imposing their legislative will on others—typically the subjects of a political community or state. By contrast, the legislation that cannot be external and does not involve the imposition of someone else’s will is “ethical” (*ethisch*) (6:218–21).

In drawing on the conceptuality of legislation and executive will for the entire sphere of practical laws, under inclusion of the specifically ethical laws of freedom, Kant extends the usage of juridico-political concepts deep into the sphere of ethics. To be sure, the originally juridical or political concepts undergo specific modifications when carried over from law to ethics. Yet they retain sufficient features from their antecedent or underlying signification to import a juridico-political basic element into Kant’s ethical discourse in the second part of the *Metaphysics of Morals*.

Most importantly, Kant extends the juridico-political notion of “constraint” (*Zwang*)—specifically, the “constraint ... of the faculty of free choice through law” (*Zwang ... der freien Willkür durchs Gesetz*)—from politics into ethics (6:379, translation modified). The term expands on the generic characterization of moral obligation involving a “necessitation” (*Nötigung*) by introducing the juridico-political term for the legal and political measures taken to assure the effectiveness of legislation. In the sphere of law and politics the common, generic feature of practical legislation (*viz.*, the law-based and law-governed con-

straint of free choice) occurs as “external constraint” (*äußerer Zwang*) or “constraint through someone else” (*Zwang durch einen Anderen*) (6:379, 396, 394, translation modified).

By contrast, the constraint involved in the sphere of ethics is internal. It does not involve someone else exercising the constraint. Rather, the constraint is exercised by the very being that undergoes the constraint. Ethical constraint is “self-constraint” (*Selbstzwang*) (6:379–81, 394, 396). More precisely, the constraint peculiar to ethical obligation or ethical duty is “free self-constraint” (*freier Selbstzwang*) (6:383, 395), as opposed to the unfree self-constraint involved in one’s faculty of free choice being constrained by a law of nature. The latter is the case when one “natural inclination” (*Naturneigung*) (6:394) endeavors to constrain another one. By contrast, in the case of free self-constraint the constraint on the faculty of free choice is not exercised by competing inclinations but through “pure practical reason” as the legislative will power that is able to constrain the influence of inclinations internally (6:396).

But ethical self-constraint is not only negatively free, due to the absence of naturally based inclinations; it can also be considered free in the positive sense insofar as it is exercised in accordance with laws of freedom that serve to realize freedom collectively by restricting it distributively. In the paradigm case of the external constraint exercised by juridical laws of freedom, legal constraint makes possible “outer freedom” (*äußere Freiheit*) (6:380, 396, 406) by limiting everyone’s freedom through the concept of everyone else’s freedom. In the parallel case of the internal constraint exercised on oneself, the constraint exercised by ethical laws renders possible “inner freedom” (*innere Freiheit*) (6:396, 405 f, 408) by limiting the influence of inclinations through the concept of the autonomy of the will. Based on the general consideration that the very possibility of intersubjective (juridical) or intrasubjective (ethical) freedom requires constraint through laws that realize freedom by restricting it, Kant measures the intensional magnitude of freedom in terms of the kind of constraint by law involved, stating: “Self-constraint is the highest degree of freedom ...” (28.1:100).

Kant marks the difference between the generic “autonomy of practical reason” (*Autonomie der praktischen Vernunft*) that holds throughout the domain of practical philosophy, insofar as reason is the ground for the validity of the laws of freedom, and the specific legislation involved in ethics by resorting to the politically inflected term, “autocracy” (*Autokratie*), to designate the latter (6:383). In the autocratic regimen of eth-

ics, the autonomous legislation of ethical laws is joined by the consciousness of one's ability or faculty (*Vermögen*) to prevail over those inclinations that run counter to the law. Kant does not claim an immediate awareness of this faculty or ability as such but grounds the mediated, inferential consciousness of it in one's immediate awareness of the ethical categorical imperative.

By resorting to the originally political concept of autocracy Kant has chosen a term that conveys the unity of legislative and executive power in one and the same being. As Kant himself puts it: in the free self-constraint involved in ethical obligation and ethical duties one's own legislative reason "constitutes itself into a power *executing* the law" (*sich zu einer das Gesetz **ausführenden** Gewalt selbst constituirt*) (6:405, translation modified). Kant designates the autocratic self-empowerment of legislative will to executive will in ethical matters with the traditional term, "virtue" (*Tugend*), meaning "moral strength of the will" (*moralische Stärke des Willens*) (6:405). He continues the use of concepts of political rule for the constitution of the ethical self, when he identifies virtue with the command "to bring all one's faculties and inclinations under one's (reason's) power" (*seine Vermögen und Neigungen unter seine [der Vernunft] Gewalt zu bringen*) or to "rule over oneself" (*Herrschaft über sich selbst*) (6:408, translation modified). In particular, Kant names a two-fold requirement for ethical autocracy: with respect to one's "affects" (*Affekten*), constituting possible distractions from ethical obligation and ethical duties through *feeling*, one is to be "master of oneself" (*seiner selbst ... **Meister***); with respect to one's "passions" (*Leidenschaften*), constituting possible distractions from ethical obligation and ethical duties through *desires*, one is to be "commander over oneself" (*über sich selbst **Herr** zu sein*) (6:407, translation modified). Drawing on the political analogy involved, one might add that the point of self-mastery and self-command in ethical matters is not the extirpation of the affects and desires but their governance, just as the point of political rule, according to classical political thought, is not the abolition of the ruled but their control.

According to Kant, the self-constraint involved in ethical conduct exhibits a further feature that attests to the particularly conflicted nature of the practical self. On the one hand, self-constraint is *freely* exercised by pure practical reason or pure will on the faculty of free choice, aimed at overriding the countervailing inclination. Due to the persisting "resistance of their inclination" (*Widerstand ihrer Neigung*), human beings, even when following the "moral law" (*moralisches Gesetz*), do so

“reluctantly” (*ungern*) (6:379, 379n, translation modified). That is exactly why there needs to be constraint on the part of the legislative will in the first place. But on the other hand, ethical conduct also involves a constraint in the opposite direction, a constraint that is exercised not against inclinations resisting ethical or free self-constraint, but in support of the purely ethical determination of the will against those inclinations that were able to prevail over free, ethical self-constraint (6:379n).

The reluctance human beings feel in obeying the commands of the moral law, if they do so, has its exact counterpart in the reluctance they feel in disobeying those very commands, if they do so. According to Kant, ethical transgression does not occur triumphantly or even contently and out of complete conviction, but reluctantly and in cognition as well as recognition of the wrongdoing involved. Hence the practical self is caught in the middle of two types of constraint that necessitate it from within and therefore represent different types of self-constraint. Moreover, the two forms of self-constraint are opposites of each other, each operating on the “faculty of free choice” (*Willkür*) in opposed directions: one constraining the resistance on the part of the inclinations against compliance with the law, the other constraining the prevailing of the inclinations toward non-compliance with the law.

Kant himself characterizes the peculiar position of the doubly reluctant practical self—a self that is reluctantly ethical as well as reluctantly unethical—as one of “mutually opposed self-constraint” (*wechselseitig entgegengesetzte[r] Selbstzwang*) and places the internally conflicted self at the “crossroads” (*Scheideweg[e]*) between “virtue and voluptitude” (*Tugend und Wohllust*) (6:379n, translation modified). But while being structurally akin, the free self-constraint and the unfree self-constraint<sup>10</sup> that in their oppositional reciprocity make up the human condition are functionally asymmetrical and do not simply cancel each other out in a relation of equipollence. Kant cites the “phenomenon” (*Phänomen*) that the human being at the crossroads shows “more propensity to listen to inclination than to the law” (*mehr Hang ... der Neigung als dem Gesetz Gehör zu geben*) (6:379n, translation modified). He here refers to an anthropological fact that he considers empirically confirmed but not susceptible of explanation. For any explanation of the phenomenon of anti-morality would involve its derivation from a cause according to laws of nature. But this naturalist explanation would contradict the assumed or presupposed freedom of our “faculty of choice” (*Willkür*),

---

10 Kant himself does not use the term, “unfree self-constraint”.

an assumption or presupposition made on strictly moral grounds and independent of the factual evidence of prevailing immorality.

Like Plato's innerly conflicted one-man *polis*, Kant's inner republic of strife faces the threat of contradiction and logical self-destruction. Moreover, like Plato, Kant resorts to an internal distinction of the self that assigns the opposed forms of self-constraints that the self both exercises and undergoes to specifically different functions of practical subjectivity. In particular, Kant attributes the reluctance to obey the moral law to the status of human beings as "rational *beings of nature*" (*vernünftige Naturwesen*), who subordinate their rationality to extra- or even irrational desires.<sup>11</sup> By contrast, he traces the reluctance to disobey the moral law to the status of human beings as "*moral beings*" (*moralische Wesen*). In the latter capacity, the human being is not subject to natural laws but must be regarded, and has to regard itself, as a "*free (moral) being*" (*freies [moralisches] Wesen*): free from outer as well as inner natural constraint and free to undergo rational self-constraint based on the "*humanity in his own person*" (*Menschheit in seiner eigenen Person*).<sup>12</sup> The term "humanity" is here to be taken not extensionally, as including all human beings, but intensionally, as excluding those traits in human beings that stem from the "*animality of the human being*" (*Thierheit des Menschen*) (6:420, translation modified).

Kant further expands on the dual composition of the practical self by contrasting the "*human animal*" (*Thiermensch*) and the "*rational human being*" (*Vernunftmensch*) (6:435). In the former capacity the human being is not just an animal but an "animal endowed with reason" (*mit Vernunft begabtes Thier*), capable of rationally informed conduct (6:456). In the latter capacity the human being is a "*rational being*" (*vernünftiges Wesen*), capable of conduct determined solely and entirely by reason (6:456).

Also like Plato, Kant resists dissolving the complex and conflicted constitution of the self into the compatibilist coexistence of distinct parts in an encompassing whole. Terms such as "humanity" and "animality", along with artificial coinages such as "human being as animal being" and "human being as rational being", manifest an attempt at a functional differentiation of the (practical) self. In line with this under-

11 On the difference between a "rational being" (*vernünftiges Wesen*) and a "being of reason" (*Vernunftwesen*), see 6:418, translation modified.

12 6:379 f. See also the related distinction between "*homo noumenon*" and "*homo phaenomenon*" in 6:418, 423.



standing Kant does not consider the presence of animality in the human being a case of partial inclusion, but resorts to the unusual prepositional locution “the animal *about* the human being” (*[das] Thier[es] am Menschen*) (6:445, translation modified). Rather than partitioning the human being, Kant undertakes a functional differentiation by attributing to the human being, in a practical regard, a “twofold personality” (*zweifache Persönlichkeit*) (6:439n, translation modified).

According to Kant, the numerically identical human being (“*numero idem*”) functions in a twofold way and is insofar different in kind (“*specie diversus*”): as a “subject of the moral legislation that originates in the concept of freedom” (*Subjekt der moralischen, von dem Begriffe der Freiheit ausgehenden Gesetzgebung*) and as a “sensory human being endowed with reason” (*der mit Vernunft begabte Sinnenmensch*) (6:439n, translation modified). The specific difference in the functioning of human practical subjectivity is attributable to the “faculties of the human being (the higher and lower ones)” (*der Fakultäten des Menschen [der oberen und unteren]*) (6:439n, translation modified). In thinking the concept of human being “not in one and the same sense” (*nicht in einem und demselben Sinn*) but as involving a dual sense of subjectivity, Kant links his moral philosophy, in particular his ethics, to the “distinction between the things as objects of experience from those same things as things in themselves” (*Unterscheidung der Dinge, als Gegenstände der Erfahrung, von eben denselben, als Dingen an sich selbst*) undertaken in the *Critique of Pure Reason* (Bxxvii). The “critical distinction” (*kritische Unterscheidung*) (Bxxviii) makes a practical reappearance as the distinction, with respect to one and the same human being, between the internally free human being (“*homo noumenon*”) and the causally determined as well as determining human being (“*homo phaenomenon*”) (6:418). As in Plato, practical philosophy in Kant draws on first philosophy, just as in Plato as well as Kant ethics draws on politics and law.

## 28. Die Person als gesetzgebendes Wesen

*Katsutoshi Kawamura*

### 1. Die Gesetzgebung der Person

In der *Grundlegung zur Metaphysik der Sitten* wird der Begriff der Person verstanden als ein „vernünftiges Wesen“ und als ein „gesetzgebendes Wesen“. Unter dem „vernünftigen Wesen“ versteht Kant ein Wesen, das unabhängig vom Naturgesetz eine nicht-sinnliche Ordnung nach eigenem Gesetz konstruieren kann, die ihrerseits „mundus intelligibilis“ genannt wird (4:438). Ein vernünftiges Wesen ist dasjenige, das „seine Maximen jederzeit aus dem Gesichtspunkte seiner selbst, zugleich aber auch jedes anderen vernünftigen als gesetzgebenden Wesen (die darum auch Personen heißen)“ (4:438), nehmen muss. In diesem Sinne ist die „Person“ mit dem „gesetzgebenden Wesen“ gleichzusetzen bzw. ist als ein Wesen zu verstehen, das sich von selbst ein solches Gesetz der Handlung gibt, das etwas anderes als Naturgesetz ist. Was bedeutet aber die „Gesetz-Gebung“, in der die Person besteht?

Nimmt man die Gesetzgebung im weiteren Sinne, so gehört u. a. die subjektiv-persönliche Regel der Handlung zu ihr. Die subjektiv-persönlichen Handlungsregeln, nach der Terminologie Kants die „Maximen“, gibt sich jeder Mensch aus eigenem Gesichtspunkt. Folglich kann man sagen, dass die Selbst-Gesetzgebung in der Maximen-Bildung besteht.<sup>1</sup> Es sind subjektive Regeln der Handlung, die als Disziplin das Alltagsleben des Betroffenen bestimmen. Sie stammen nicht nur aus den von der Vernunft gesteuerten Überlegungen, sondern auch und in erster Linie aus der subjektiv-persönlichen Eigenschaft, d. h. nach der Terminologie Kants, aus der Neigung. Diese Regeln des Alltagslebens sind

---

<sup>1</sup> Beatrix Himmelmann z. B. erläutert, dass die Maxime aus dem freien Willen des Subjekts stammt und folglich als Autonomie zu verstehen ist: „Ohne Frage haben die Maximen oder subjektiven Prinzipien meines Handelns ihren Ursprung in meinem freien Willen und sind in diesem Sinn Ausdruck meiner Autonomie“ (Beatrix Himmelmann, *Kants Begriff des Glücks* [Berlin, New York: Walter de Gruyter, 2003], 54).

empirisch nachprüfbar, und ihr Wesen besteht gerade in der eigenen Gesetzgebung. Sie können und müssen jedoch nicht allgemeine Gültigkeit besitzen. Die eigene Gesetzgebung subsumiert hier in erster Linie deren subjektiv-persönliche Gültigkeit unter sich. Kants Beispiel lautet: Ich will zu keines Menschen Wohlbefinden oder Beistand in der Not etwas beitragen (vgl. 4:423).<sup>2</sup> Oder: Die anderen Menschen müssen mir nicht wohl tun, wenn ich es überhoben sein dürfte, ihnen Wohltat zu erzeugen (vgl. 4:430 Anm.). Die Gemeinsamkeit dieser Maximen heißt etwa, dass man im Leben möglichst wenig Umgang mit den anderen Menschen haben möchte. In diesen Maximen sieht man die Ablehnung der Reziprozität unter den Menschen, welche die Grundlage des gesellschaftlichen Lebens zu verstehen ist und die gerade in der als Zweckformel des kategorischen Imperativs<sup>3</sup> erfordert wird.

Insoweit die eigene Gesetzgebung aus der eigenen Neigung stammt, folgt man ihr von der eigenen Natur her, und es besteht dabei kein innerliches Hindernis. Die Selbst-Gesetzgebung und deren Befolgung stimmen hier im Prinzip ohne Hindernis spontan überein, d. h. die Gesetzgebung bedeutet zugleich deren Befolgung. Hier sieht man das Muster der Selbst-Gesetzgebung, die jeder sich selbst tatsächlich gibt und durch die er seinen eigenen Charakter äußert. Um die Gesetzgebung als solche festzustellen, muss man von deren Ergebnis, d. h. von der Maxime zurückgehen. Sie als solche ist im Prinzip immer schon als vergangen zu verstehen und kann nur retrospektiv festgestellt werden.

Betrachtet man die „Gesetz-Gebung“ im engeren Sinne, so erkennt man, dass das Gesetz nicht subjektiv-individuell sein darf, sondern objektiv-allgemeingültig sein muss, insofern es ein moralisches Gesetz ist. Kant erläutert, „Jedermann muß eingestehen, daß ein Gesetz, wenn es moralisch, d. i. als Grund einer Verbindlichkeit gelten soll, absolute Notwendigkeit bei sich führen müsse; daß das Gebot: Du sollst nicht lügen, nicht etwa bloß für Menschen gelte, andere vernünftige Wesen sich daran nicht zu kehren hätten; und so alle übrigen eigentlichen Sittengesetze“ (4:389). In diesem Sinne kann ein Gesetz nur dann gegeben werden, wenn der Gesetzgeber von den subjektiv-individuellen Bedin-

---

2 Vgl. Maria Schwarz, *Der Begriff der Maxime bei Kant. Eine Untersuchung des Maximenbegriffs in Kants praktischer Philosophie* (Berlin: LIT Verlag, 2006), 158 (Anhang).

3 „Handle so, daß du die Menschheit, sowohl in deiner Person, als in der Person eines jeden anderen, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchst“ (4:429).

gungen unabhängig sein kann, oder wenn er sie gegebenenfalls ablehnen kann. Diese Unabhängigkeit bzw. Ablehnbarkeit setzt ihrerseits die Freiheit voraus, die eine Handlung oder ein Zustand von selbst anzufangen bedeutet. Diese Freiheit wird von Kant zunächst in der *Kritik der reinen Vernunft* in der Antinomienlehre abgehandelt. Sie wird als die absolute Spontaneität des handelnden Subjekts verstanden, die als eine Idee nur außerhalb der Sinnenwelt, in der alles nur nach dem Naturgesetz abläuft, gedacht werden kann.<sup>4</sup> Die Beweisbarkeit dieser Freiheit lehnt Kant in der *ersten Kritik* eindeutig ab (vgl. B585 f.). Diese Freiheit, d. h. die transzendente Freiheit, wird aber in der *Kritik der praktischen Vernunft* erneut festgestellt, und zwar als der Seinsgrund des moralischen Gesetzes (vgl. A5 Anm.). Dieser Seinsgrund, der als absolute d. h. nicht bedingte Spontaneität zu verstehen ist, ist im Kontext der praktischen Reflexion die Basis der Selbst-Gesetzgebung, die gemäß der Tätigkeit der absoluten Spontaneität ebenfalls außerhalb des sinnlich-nachprüfbaren Bereichs gedacht werden muss. Hier zeigt sich die sogenannte Zweiweltentheorie Kants deutlich, nach der mundus sensibilis und mundus intelligibilis sowie Phaenomenon und Noumenon in ein- und demselben Menschen unterschieden werden müssen.

Diese Selbst-Gesetzgebung wird in der *Grundlegung* folgendermaßen formuliert: „Autonomie des Willens ist die Beschaffenheit des Willens, dadurch derselbe ihm selbst (unabhängig von aller Beschaffenheit der Gegenstände des Willens) ein Gesetz ist. Das Prinzip der Autonomie ist also: nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Willen zugleich als allgemeines Gesetz mit begriffen seien.“ (4:440). Diese Unabhängigkeit des Willens von Beschaffenheit der Willens-Gegenstände bedeutet zugleich die Unabhängigkeit vom Naturgesetz, welches durch die Menschennatur via Begierde, Instinkt oder Interesse den Willen bestimmt, und bedeutet ebenfalls die o.g. Freiheit des handelnden Subjekts. Aus dem Ausdruck, dass das Gesetz die Beschaffenheit des Willens ist, lässt sich vermuten, dass dieses Gesetz bereits im Willen virtuell vorhanden ist.

---

4 Vgl. „... Dagegen verstehe ich unter Freiheit, im kosmologischen Verstande, das Vermögen, einen Zustand von selbst anzufangen, deren Kausalität also nicht nach dem Naturgesetze wiederum unter einer anderen Ursache steht, welche sie der Zeit nach bestimmte“ (B561). Auch Kawamura, *Spontaneität und Willkür. Der Freiheitsbegriff in Kants Antinomienlehre und seine historischen Wurzeln* (Stuttgart-Bad Cannstatt: Frommann Holzboog, 1996), insbes. 137–78.

Hier fragt man sich, ob die Selbst-Gesetzgebung als Beschaffenheit des Willens immer schon im Willen vorhanden ist, oder ob sie als eine Tätigkeit irgendwann geschaffen werden muss.

## 2. Erste These

Eine Antwort kann lauten: die Selbst-Gesetzgebung ist die Beschaffenheit des Willens, die zeitunabhängig immer schon im Willen vorhanden ist, weil diese Tätigkeit der Gesetz-Gebung sich auf die transzendente Freiheit bzw. die absolute Spontaneität gründet, welche ihrerseits außerhalb der Sinnenwelt nur als eine Idee der Selbsttätigkeit des handelnden Subjekts gedacht werden kann. Diese Selbst-Gesetzgebung und deren Ergebnis, nämlich das moralische Gesetz, lassen sich als Bedingung aller einzelnen Werturteile als immer schon vorhanden denken. Sie mag wohl eine Vorgabe der reinen praktischen Vernunft sein, die ihrerseits als vor allen möglichen Werturteilen a priori vorhanden zu verstehen ist. Hier erinnert man sich an die Selbsttätigkeit des Erkenntnissubjekts im Bereich der theoretischen Vernunft, vor allem deren Konstitution der möglichen Erfahrung überhaupt. Das Erkenntnissubjekt handelt außerhalb der raum-zeitlichen Dimension bzw. außerhalb der Sinnenwelt, und zwar vor dem Entstehen dieser Dimension. Dieses Subjekt beschäftigt sich mit der Konstitution der Bedingung aller möglichen Erfahrungen bzw. aller möglichen Gegenstände überhaupt, ohne dessen Tätigkeit die Erfahrungen und deren Summe als die Sinnenwelt nicht zustande kommen. Es lässt sich vermuten, dass die Selbst-Gesetzgebung des Willens als die Tätigkeit der reinen praktischen Vernunft, genauso wie die Selbsttätigkeit des Erkenntnissubjekts, vor aller möglichen Erfahrung bzw. vor allen möglichen Werturteilen als deren Bedingung zeitunabhängig vorausgesetzt zu denken ist. Die Autonomie des Willens ist also nicht als die Aufgabe, sondern als die Vorgabe der reinen praktischen Vernunft zu denken. In *Erste Einleitung in die Kritik der Urteilskraft* lässt sich erkennen, dass Kant unter der „Autonomie“ die apriorisch gesetzgebende Tätigkeit überhaupt versteht. D.h. nicht nur die Selbst-Gesetzgebung der praktischen Vernunft, sondern auch die des Verstandes wird als Autonomie verstanden.<sup>5</sup>

---

5 In der von Gerhard Lehmann herausgegebenen *Erste Einleitung in die Kritik der Urteilskraft* liest man die folgende Sätze: die Urteilskraft „die ... in Ansehung der Bedingungen der Reflexion a priori gesetzgebend ist und Autonomie beweiset;

Es entspricht der Auffassung Kants, dass man immer schon weiß, was das moralische Gesetz ist.<sup>6</sup> Mit dem moralischen Gesetz, das jeder immer schon weiß, sind m. E. die allgemein anerkannten moralischen Gebote gemeint, wie z. B. „du sollst nicht lügen“ (4:389), die Kant selbst als Beispiel der moralischen Gesetze anführt, oder „du sollst nicht töten“ u. a., deren Herkunft sich in der europäischen Kulturtradition bis auf das Alte Testament zurückverfolgen lässt. Oder es kann damit auch die Handlungsregel gemeint sein: „Was du nicht willst, das man dir tut, tue auch keinem anderen“ (vgl. 4:430 Anm.), die ebenfalls spätestens seit der Zeit des Alten Testaments vorhanden ist, und zwar nicht nur unter bestimmten Völkern bzw. Religionskreisen, sondern unabhängig von den Religionen und Kulturen.<sup>7</sup> (Obwohl Kant in der *Grundlegung* die Gültigkeit dieser Regel, nämlich der „Goldenen Regel“, als Prinzip der Moral kritisiert, lässt sich deren allgemeine Gültigkeit als Moralprinzip und die Gemeinsamkeit mit dem kategorischen Imperativ nicht leugnen<sup>8</sup>). Als dasjenige moralische Gesetz, das jeder kennt, kann man die hier erwähnten Gebote oder Regeln betrachten, die ohne Zweifel eine allgemeine Gültigkeit besitzen und deren Sinn jeder Mensch versteht. In der *Grundlegung* wird erläutert, dass „das allgemeine Prinzip der Sittlichkeit ... in der Idee allen Handlungen vernünftiger Wesen ebenso zum Grunde liegt, als das Naturgesetz allen Erscheinungen“ zum Grunde liegt (4:452 f.). Aus dieser parallelen Erklärung lässt sich erkennen, dass Kant

---

diese Autonomie aber ist nicht (so wie die des Verstandes in Ansehung der theoretischen Gesetze der Natur, oder der Vernunft in praktischen Gesetzen der Freiheit) objektiv ...“ (20:225); Nach der Handschrift herausgegeben von Gerhard Lehmann, 4. Aufl. (Hamburg: Felix Meiner, 1990), 32.

- 6 In der *KpV* im Zusammenhang mit dem Begriff der Freiheit erläutert Kant: „Freiheit ist ... auch die einzige unter allen Ideen der spekulativen Vernunft, wovon wir die Möglichkeit a priori wissen, ohne sie doch einzusehen, weil sie die Bedingung des moralischen Gesetzes ist, welches wir wissen“ (A5).
- 7 Inhaltlich der gleiche Satz findet sich in der alten chinesischen Schrift „Lunyü“ (Gespräch) von Konfuzius, vgl. Konfuzius, *Gespräche (Lunyü)*, in der Übersetzung von R. Wilhelm. Neu gesetzte und überarbeitete Ausgabe (Wiesbaden: Marix Verlag, 2005), Buch XII 2, 173: „Was du selbst nicht wünschst, das tue nicht den Menschen an“; Buch XV 23, 231: „Die Nächstenliebe. Was du selbst nicht wünschst, tu nicht an andern“.
- 8 Vgl. H. H. Schrei u. H. U. Hoche, Artikel „Regel, goldene“ in: *Historisches Wörterbuch der Philosophie*, hrsg. Von J. Ritter, Bd. 8 (Basel: Schwabe & Co., 1987), Sp.450–64. Kawamura, „Kants Kritik an der Goldenen Regel“, in *Kant zwischen West und Ost. Zum Gedenken an Kants 200. Todestag und 280. Geburtstag*, hrsg. W. Bryuschinkin 2. Bde. (Kaliningrad: Kaliningrad University Press, 2005), Bd. 2, 179–86, insbes. 179.

unter dem allgemeinen Prinzip der Sittlichkeit, vermutlich dem moralischen Gesetz im engeren Sinne, so etwas wie das Naturgesetz versteht, welches für alle Geschehnisse allgemeine Gültigkeit besitzt.

Dass die Gesetzgebung des Willens als a priori gegeben aufzufassen ist, entspricht ebenfalls der Äußerung Kants betreffend des guten Willens, dass jeder weiß, was ein guter Wille ist, ohne jede Erziehung oder Bildung.<sup>9</sup> Ein guter Wille ist derjenige, der sich ständig nach dem moralischen Gesetz, und zwar um des moralischen Gesetzes willen, zur Handlung bestimmt. In der praktischen Philosophie Kants hängen das moralische Gesetz und der gute Wille wesentlich zusammen. Insoweit das moralische Gesetz immer schon vorhanden ist, muss dessen Gesetzgebung ebenfalls immer schon vorhanden und a priori sein.

### 3. Zweite These

Jedoch, wie soll diese apriorische Gesetz-Gebung mit der Tatsache übereinstimmen, dass man sich u. U. des moralischen Gesetzes überhaupt nicht bewusst ist, oder dass es wenigstens so aussieht? Falls die Autonomie eine Art der Vorgabe ist, aus welchem Grund kann man sie dann so häufig ignorieren oder ihr widersprechend handeln? Diesen Fragen liegt die oben genannte Frage zu Grunde, ob die Selbst-Gesetzgebung als eine Tätigkeit irgendwann geschaffen wurde.

Eine Antwort könnte lauten: diese Gesetz-Gebung muss irgendwann im Prozess der Entwicklung eines jeden handelnden Subjekts durchgeführt werden, und sie muss ebenfalls bei wichtigen Entscheidungen betreffend des Werturteils erneut aktiviert werden, sonst bliebe sie bloß eine potenzielle Beschaffenheit des Willens. Sieht man z.B. in der wirtschaftlichen Krise viele Obdachlose, so fragt man sich, ob man bei der Maxime bleiben kann: Ich will zu keines Menschen Wohlbefinden oder Beistand in der Not etwas beitragen. Bei dieser Fragestellung versteht man wenigstens, dass diese Maxime nicht verallgemeinerbar ist, weil Bedürftige diese Maxime vermutlich nicht wollen können. Durch dieses

---

9 Die Überschrift des ersten Abschnitts der *Grundlegung* lautet „Übergang von der gemeinen sittlichen Vernunftkenntnis zur philosophischen“ (4:393). Gleich nach dieser Überschrift wird erläutert, dass ein guter Wille das einzige sei, was „ohne Einschränkung für gut könnte gehalten werden“. Es lässt sich vermuten, dass Kant unter der „gemeinen sittlichen Vernunftkenntnis“ u. a. unsere Erkenntnis vom guten Willen versteht.

Verständnis kann man die eigene Maxime korrigieren und zwar gemäß dem Maßstab der Verallgemeinerbarkeit. Jedoch scheint nicht jede Maxime verallgemeinerbar zu sein. Bei jeder Maximenbildung entsteht die Diskrepanz zwischen der gegebenen Maxime und dem gedachten allgemeinen Gesetz. Anders formuliert, lässt sich das moralische Gesetz als solches nicht eindeutig begreifen. Nach Kant soll das moralische Gesetz allgemeiner sein als die goldene Regel (4:430 Anm.). Das moralische Gesetz im engeren Sinne, d. h. im Singular, ist als die von der goldenen Regel oder auch von den o.g. Befehlen, wie „du sollst nicht töten“, abgeleitete Idee eines allgemein-gültigen Gesetzes zu verstehen, die Kant niemals als solche zum Ausdruck gebracht hat.

Die zweite Frage bezieht sich auf die Diskrepanz zwischen der Selbst-Gesetzgebung und deren Befolgung. Falls das moralische Gesetz als von Anderen gegeben zu verstehen ist, so ist eine diesem Gesetz widerstrebende Handlung nicht selbst-widersprüchlich. Insoweit es sich jedoch um Selbst-Gesetzgebung handelt, ist es ersichtlich, dass das Subjekt dem selbstgegebenen Gesetz unterworfen ist und ihm nicht folgen kann. Wie oben gesehen, stimmen bei der Selbst-Gesetzgebung der subjektiven Handlungsregeln bzw. der Maxime Selbst-Gesetzgebung und deren Befolgung innerlich überein.

Falls das moralische Gesetz den Handelnden nicht zureichend motivieren kann, fragt man sich, was ihn dann eigentlich zur Handlung motivieren könne. Nach Kant mag es ein spezifisches Gefühl sein, welches moralisch genannt wird. Es „ist ein Gefühl, was bloß aufs Praktische geht und zwar der Vorstellung eines Gesetzes lediglich seiner Form nach, nicht irgend eines Objekts desselben wegen anhängt, mithin weder zum Vergnügen noch zum Schmerz gerechnet werden kann und dennoch ein Interesse an der Befolgung desselben hervorbringt, welches wir das moralische nennen“ (A142). Es mag ein spezifisch menschliches (und nicht tierisches) Gefühl sein, das einen motiviert, gemäß dem moralischen Gesetz zu handeln. Kant sieht in der Reflexion des Menschen über sich selbst eine „Achtung fürs moralische Gesetz“ (A142), und diese Achtung äußert sich in der Möglichkeit, die Klugheitsregel abzulehnen und nach dem moralischen Gesetz zu handeln. Diese Achtung scheint ihrerseits ein Zugang zu sein, durch welchen man sich das moralische Gesetz als solches deutlicher vorstellen kann.

In diesem Zusammenhang erinnert man sich ebenfalls an das Begriffspaar „Sinnenwelt“ und „Verstandeswelt“ (4:451) in ein- und demselben Menschen. Nach Kant ist der Mensch als Zugehöriger zur Verstandeswelt das „Selbst“ der Selbst-Gesetzgebung, und als Zugehö-



riger zur Sinnenwelt ist er dem Gesetz unterworfen. Als der Erstere ist er das eigentliche Subjekt der Autonomie und zugleich das der absoluten Spontaneität bzw. der Freiheit, während er als Letzterer dem gegebenen moralischen Gesetz unterworfen ist und als darunter leidend zu verstehen ist.<sup>10</sup> In der *Kritik der praktischen Vernunft*, nachdem Kant den Begriff der Persönlichkeit als „die Freiheit und Unabhängigkeit von dem Mechanismus der ganzen Natur“, d. h. als zur Verstandeswelt gehörig definiert hat, erläutert er, dass „die Person also als zur Sinnenwelt gehörig ihrer eigenen Persönlichkeit unterworfen ist, sofern sie zugleich zur intelligibelen Welt gehört“ (A155). Während hier die Persönlichkeit eindeutig als zum mundus intelligibilis gehörig aufgefasst wird, wird die Person als zu den beiden Welten gehörend verstanden. Die Vorstellung von homo noumenon kann durch die o.g. Achtung fürs moralische Gesetz ebenfalls verdeutlicht werden.

#### 4. Schlussbemerkung

Aus diesen Überlegungen lässt sich schließen, dass die Selbst-Gesetzgebung der Person in ihrer Tätigkeit wenigstens zwei Seiten hat, die wechselseitig unter einer zirkelhaften Beziehung stehen. Einerseits ist die Tätigkeit der Person auf der Ebene des Menschen als homo noumenon vor und außerhalb der möglichen Erfahrung überhaupt von selbst gesetzgebend, und deren Ergebnis, das moralische Gesetz, ist als ein Kriterium des Guten a priori als immer schon vorhanden zu verstehen. Die Gesetz-Gebung ist in diesem Sinne als Vorgabe für uns Menschen zu verstehen. Andererseits bedeutet die Selbst-Gesetzgebung für uns eine Aufgabe, welche ein dem immer schon gegebenen moralischen Gesetz unterworfenen „Selbst“ voraussetzt und dass die vorhandene eigene Maxime gemäß der Vorstellung eines allgemeingültigen Gesetzes zu korrigieren ist. Die Gesetz-Gebung von dem moralischen Gesetz im Singular kann in diesem Zusammenhang als die ideelle Tätigkeit nur von deren Ergebnis, dem moralischen Gesetz her rückwirkend gedacht werden. Die Realität dieser Gesetz-Gebung hängt ständig von der des moralischen Gesetzes ab, dessen Realität nur durch die Gesetz-Gebung

---

10 Vgl. 4:440: Dass „wir uns dadurch zugleich eine gewisse Erhabenheit und Würde an derjenigen Person vorstellen, die alle ihre Pflichten erfüllt. Denn sofern ist zwar keine Erhabenheit an ihr, als sie dem moralischen Gesetz unterworfen ist, wohl aber sofern sie in Ansehung ebendesselben zugleich gesetzgebend und nur darum ihm untergeordnet ist“.

im Sinne der Korrektur der eigenen Maxime festgestellt werden kann. „Unterworfen sein“ heisst hier, dass die Person sich jederzeit schon unter dem moralischen Gesetz befindet und dass eben deshalb die Gesetzgebung als Maximen-Bildung bzw. Korrektur der Maxime für die Person als homo phaenomenon immer als Aufgabe zu verstehen ist. Hier sieht man die zwei Dimensionen der Person, die zusammen einen Zirkel ausmachen, in welchem eine allgemeingültig gesetzgebende, sowie eine dem Gesetz unterworfenen und nach dem Gesetz strebende Person unterschieden werden.

## 29. Kant's Realm of Ends: A Communal Moral Practice as Locus for the Unity of Moral Personhood

Stijn Van Impe

### 1. A Reevaluation of Kant's Moral World

Throughout his mature writings Kant offers several assessments of a moral world (i. e., a world not as it is through mechanical laws of nature, but as it should be through moral laws of freedom). He approaches this idea alternatively in terms of a “realm of grace” in the *Critique of Pure Reason* (CPR), a “realm of ends” in the *Groundwork of the Metaphysics of Morals* (GMM) and *Lectures on the Philosophical Doctrine of Religion* (LPR), a “kingdom of God” in the *Critique of Practical Reason* (CPrR), and a “realm of virtue” or “ethical commonwealth” in *Religion within the Boundaries of Mere Reason* (RBR). The importance of this topic for Kant's moral philosophy can thus hardly be overestimated, especially since Kant strikingly holds that “Morality consists, then, in the reference of all action to the lawgiving by which alone a realm of ends is possible” (4:434) and that “Morality ... is an absolutely necessary system of *all ends*, and it is just this agreement with the idea of a system which is *the ground of the morality of an action*” (28:1075). Despite the vast amount of literature on Kant's moral philosophy in general, and GMM in particular, little attention has been given to Kant's realm of ends.<sup>1</sup>

This state of affairs yields the following pressing questions. First, what exactly does Kant understand by the realm of ends? Does it

---

1 See Kurt Bache, “Kants Prinzip der Autonomie im Verhältnis zur Idee des Reiches der Zwecke”, *Kant-Studien Ergänzungsheft* 12 (Berlin: Walter de Gruyter, 1909); Mary A. McCloskey, “Kant's Kingdom of Ends”, *Philosophy* 51 [198] (1976), 391–9; Christine M. Korsgaard, *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996); Barbara Herman, “A Cosmopolitan Kingdom of Ends”, in Andrews Reath, et al. (eds.), *Reclaiming the History of Ethics: Essays for John Rawls* (New York: Cambridge University Press, 1997), 187–213; and Sarah Holtman, “Autonomy and the Kingdom of Ends”, in Thomas E. Hill Jr. (ed.), *The Blackwell Guide to Kant's Ethics* (Chichester: Wiley-Blackwell, 2009), 102–17.

have a multi-layered meaning or is its meaning quite straightforward? Secondly, what does Kant mean when he calls the realm of ends alternatively an ideal and an idea of pure practical reason? To my knowledge, insufficient attention has been given to applying Kant's understanding of the notion of an ideal in the first *Critique* to the realm of ends from *GMM*. Thirdly, what does Kant have in mind when he refers to the realm of ends in terms of a "complete determination" of the categorical imperative? I believe that the significance of the realm of ends as complete determination of the categorical imperative has not yet received the valuation it truly deserves especially as it stresses the social and communal dimension of Kant's ethics.

This essay will therefore have the following set-up. First, I will examine Kant's definition of the realm of ends from the *GMM* and Kant's lesser known *LPR*. I will argue that it has a multi-layered meaning since it refers to the union of rational lawgiving beings as ends in themselves and the unification of their morally legitimate subjective ends as well as of their guiding maxims that moral actions ought to be based upon. Secondly, I will analyze what Kant means by relating the realm of ends to the "complete determination" of the categorical imperative (i. e., "totality" combining "unity" and "plurality" in the form and matter of the will, respectively). Thirdly, I will discuss how the realm of ends functions as an "ideal" of practical reason or as a normative "archetype" grounding a truly communal moral practice. Throughout these sections, I will emphasize the social and communal aspects and dimensions of Kant's realm of ends, often neglected, downplayed, or denied.<sup>2</sup>

## 2. The Union of Rational Beings as Ends in Themselves

In *GMM* Kant defines the realm of ends (*Reich der Zwecke*) as "a whole of all ends in systematic connection (a whole both of rational beings as ends in themselves and the ends of his own that each may set himself)", and adds that "all rational beings stand under the *law* that each of them is to treat himself and all others *never merely as means* but always *at the same*

---

2 As Jennifer Mills Moore, "Kant's Ethical Community", *Journal of Value Inquiry* 26 (1992), 51, aptly points out, "Kant, his critics claim, fails to do justice to the social and interpersonal dimension of ethics. He is incurably atomistic; even his 'kingdom of ends' is nothing more than an aloof harmony of separate, individual wills."

*time as ends in themselves*". This leads to "a systematic union of rational beings through common objective laws, that is, a kingdom, which can be called a kingdom of ends (admittedly only an ideal) because what these laws have as their purpose is just the relation of these beings to one another as ends and means" (4:433).

First of all, the realm of ends entails the systematic union of rational beings who treat themselves and each other never merely as means, but always as ends in themselves. Or, as Kant writes in *LPR*, the realm of ends is "a *practical* system of ends, i. e., a system in accordance with the laws of free volition" where "every rational creature is combined with every other as reciprocal end and means" (28:1102–3). This union of rational beings as ends in themselves is a "systematic" union because it will only be established if all rational beings live up to the requirements of the categorical imperative, especially in its formula of humanity: "so act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" (4:429). This union is thus systematic—not merely contingent—because it is brought about by respecting universally valid and "objective standards that are applicable to each by virtue of [the] rationality" of each rational being as end in itself.<sup>3</sup>

Two questions immediately arise here. First, what does Kant understand by rational beings: what is it that makes a being a rational being? Secondly, what does Kant mean by treating a rational being as an "objective end" or an "end in itself"? Regarding the first question, Kant argues in *GMM* that "Rational nature is distinguished from the rest of nature by this, that it sets itself an end" (4:437). What seems to be characteristic for a rational being is thus that it has the capacity of setting and hence choosing ends for itself. Keeping in mind that it is a task of the rational being to respect humanity in oneself and others, we can turn to the *Metaphysics of Morals* (*MM*), where Kant contrasts "humanity" to "animality" and characterizes it as that aspect of our nature "by which alone" the human being "is capable of setting himself ends" (6:387). Likewise, a few pages later Kant writes (6:392): "The capacity to set oneself an end—any end whatsoever—is what characterizes humanity (as distinguished from animality)." Guyer is therefore right in arguing that

---

3 Holtman, 107.

rational being ... is the ability to set one's own ends rather than have them determined by anything other than one's own choice .... Thus to make humanity or rational being whether in oneself or anyone else the object ... means to treat the capacity of every human being as a rational being ... *to make his own choices of ends* as the sole thing of unconditional value, not to be restricted or compromised by any particular pursuit of all particular ends.<sup>4</sup>

Regarding the second question, Kant gives the following instructive account of what an "objective end" or an "end in itself" is (4:428):

Now I say that the human being and in general every rational being *exists* as an end in itself, *not merely as means* to be used by this or that will at its discretion; instead he must in all his actions, whether directed to himself or also to other rational beings, always be regarded *at the same time as an end*. ... [R]ational beings are called *persons* because their nature already marks them out as an end in itself, that is, as something that may not be used merely as a means, and hence so far limits all choice (and is an object of respect). These, therefore, are not merely subjective ends, the existence of which as an effect of our action has a worth *for us*, but rather *objective ends*, that is, beings the existence of which is in itself an end, and indeed one such that no other end, to which they would serve *merely as means*, can be put in its place, since without it nothing of *absolute worth* would be found anywhere.

As McCloskey makes clear, an end is to be understood as an "object of value" and what differentiates an objective end or an end in itself from a subjective end is that the former is "an object of value because it is adopted as a goal by an individual subject or agent and one whose value does not depend upon being so adopted".<sup>5</sup> In *GMM* Kant thus contrasts objective ends or ends in themselves with subjective ends or ends to be produced, as objects, results, or states of affairs we pursue in our actions for the sake of bringing them about. In accordance with the passage quoted above, only "people and not projects or states of mind are 'objective ends', and their being objective rather than subjective ends entails that persons are objects of value no matter how any individual subject (agent) purposes".<sup>6</sup> Persons have an absolute value (i. e., a worth or a dignity), for their status cannot be reduced to a (market or fancy) price that has only a relative value and can hence be compared and traded: "What has a price can be replaced by something else

4 Paul Guyer, *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide* (London: Continuum, 2007), 92.

5 McCloskey, 393.

6 McCloskey, 395

as its *equivalent*; what on the other hand is raised above all price and therefore admits of no equivalent has a dignity” (4:434). Or, as Kant explains in *MM*, “dignity” is to be understood as “a worth that has no price, no equivalent for which the object evaluated could be exchanged” (6:462). In *LPR*, Kant argues that rational creatures have this worth *only insofar as* they can be regarded as *members* of a realm of ends: “morality, through which a system of all ends is possible, gives to the rational creature a worth in and for itself by making it a member of this great realm of all ends” (28:1099–100; cf. 28:1204). Or, as he writes in *GMM*: “morality is the condition under which alone a rational being can be an end in itself”; it is the condition for a rational being to have worth or dignity, “since only through this is it possible to be a law-giving member in the kingdom of ends” (4:435). This adds an important *communal* dimension to Kant’s view of morality: we can gain worth only by being a member in a realm of ends; we cannot gain this worth in a condition of atomistic isolation.

The realm of ends conceived as the systematic union of rational law-giving beings who are to be viewed as objective ends, ends in themselves or persons who have an incomparable worth that stands on itself, and who are never to be treated as mere means to realizing a relative or subjective end (i. e., a purpose or a goal adopted by and aimed at by another rational being) also appeals to Kant’s definition of “organization” in the third *Critique*: “For in such a whole each member should certainly be not merely a means, but at the same time also an end, and, insofar as it contributes to the possibility of the whole, its position and function should also be determined by the idea of the whole” (5:375n). In case of the realm of ends, this idea of the whole concerns the idea of a community of rational beings who regard themselves and each other as persons (i. e., not merely as a means to an end, but at the same time also as ends in themselves possessing incomparable worth). The idea of the whole thus also determines their position and function: all persons are *members* of equal worth in the realm of ends, meaning they have an *equal share* in giving universal moral laws in this realm in such a way that they are both *lawgiver of* and *subject to* these laws. As Kant argues, “it is nothing less than the *share* it affords a rational being *in the giving of universal laws* [that] makes him fit to be a member of a possible kingdom of ends, which he was already destined to be by his own nature as an end in itself” (4:435).

### 3. The Unification of Morally Legitimate Ends

However, the realm of ends is not merely the union of all rational beings as ends in themselves (i. e., as persons having an incomparable, absolute, and inner worth instead of a merely comparative price), but also of the morally legitimate ends that each rational being may pursue. Remember that Kant speaks of the realm of ends in terms of “a whole both of rational beings as ends in themselves *and* the ends of his own that each may set himself” (4:433, italics added). This double sense of the realm of ends is crucial, but is sometimes understated or misconceived even by influential authors such as Rawls, Baur, Melnick, and Hare.<sup>7</sup> Contra these authors, I concur with Höffe and Guyer<sup>8</sup> that the realm of ends has a twofold sense, for it concerns not only the systematic connection of all rational beings (i. e., ends in themselves or objective ends), but also the systematic connection of all their particular, freely chosen subjective ends. In this respect, McCloskey accurately refers to the realm of ends as the harmonization of both objective *and* subjective ends.<sup>9</sup> By the latter Kant refers to those ends that can be called morally legitimate or permissible (i. e., ends that are directly entailed by the moral law, or ends that are at least not contradictory to or inconsistent with the moral law). In the broader sense, the realm of ends does not by definition exclude the striving for non-moral ends as long as these do not interfere or

---

7 Rawls, Baur, Melnick and Arp restrict the realm of ends to the union of rational lawgiving beings but omit the second facet of Kant's realm of ends (i. e., the unification of all morally legitimate ends that these rational beings set). See John Rawls, *Lectures on the History of Moral Philosophy* (Cambridge: Harvard University Press, 2000), 204; Michael Baur, “Kant's ‘Moral Proof’: Defense and Interpretation”, *Proceedings of the American Catholic Philosophical Association* (2000), 150; and Arthur Melnick, “Kant's Formulations of the Categorical Imperative”, *Kant-Studien* 93 (2002), 302–3; Contrary to these authors, R. M. Hare, *Freedom and Reason* (Oxford: Clarendon Press, 1963), 179, interprets the realm of ends as “the liberal's ideal” of “a good society” wherein the “ends and interests of all are given equal consideration”, and thus acknowledges that in the realm of ends the ends of all are to be taken into account, but nonetheless downplays the necessity of a *systematicity* and *totality* of all morally legitimate ends.

8 Otfried Höffe, *Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar* (Frankfurt am Main: Vittorio Klostermann, 2000), 186, and Guyer, *Kant's Groundwork*, 99–100. See also Dieter Schönecker and Allen Wood, *Kants Grundlegung zur Metaphysik der Sitten* (Schöningh: UTB Verlag, 2007), 160, 163.

9 McCloskey, 398–9.



conflict with the moral law. In the stricter sense, the ends involved are those that *necessarily* belong to the realm of ends because they can be approved and pursued by a community of rational self-legislators.<sup>10</sup> Taking these approaches together, the realm of ends requires both “the *exclusion* of ends that *in principle* cannot be shared between rational beings” and “the *furthering* of ends that unite people (such as those involving mutual respect and mutual aid).<sup>11</sup> Rational beings can therefore form a realm of ends only insofar as their ends harmonize into a system (i. e., insofar as these ends are both mutually consistent and reciprocally supportive): “a ‘realm’ requires a harmony ... of ends so that the ends of all can be pursued in common”.<sup>12</sup>

The claim that the realm of ends entails a harmonious unification or a consistent whole of morally legitimate ends to be pursued by a community of rational beings is evidenced by Kant’s statement from *LPR* where he strikingly contends: “For it is only insofar as *all* rational creatures act according to these eternal laws of reason that they stand under a principle of *community* and *together* constitute a *system of all ends*” (28:1100, italics added). Moreover, in *GMM* (4:430) Kant argues that

humanity might indeed subsist if no one contributed to the happiness of others but yet did not intentionally withdraw anything from it; but there is still only a negative and not a positive agreement with *humanity as an end in itself* unless everyone also tries, as far as he can, to further the ends of others. For, the ends of a subject who is an end in itself must as far as possible be also *my* ends.

Guyer argues that the “general laws” of the realm of ends require that each member is to assist other members in pursuing their ends—insofar as these ends are compatible with these laws—because “to treat each person as an end in himself ... is to treat each person’s ability to set her own ends freely and pursue them effectively as an end in itself and a limiting condition on all other choices”. From this he concludes:

---

10 Bruce Aune, *Kant’s Theory of Morals* (Princeton: Princeton University Press, 1979), 109, nicely argues that the realm of ends is to be conceived as “a world in which each person pursues his private ends without thereby interfering with others’ pursuit of their private ends, and in which ‘everyone endeavors also, so far as in him lies, to further the ends of others’” and concludes that “A world of this kind involves a ‘harmony’ of human ends.” See also Moore, 55, 60.

11 Allen Wood, *Kant’s Ethical Thought*, (New York: Cambridge University Press, 1999), 169.

12 Wood, *Kant’s Ethical Thought*, 166.

the first part of the concept of a realm of ends—that is, the requirement to establish a whole of rational beings as ends in themselves—entails the second part—that is, the requirement to promote a systematic and coherent whole of the particular ends that each may set for herself.<sup>13</sup>

Yet, if the ends of other persons are to be my ends as well, and if, vice versa, the ends I as a person set are to be the ends of other persons as well, it becomes evident that all these ends have to be pursued *in common* (i. e., all persons have to assist—as far as possible—in realizing each other's ends). Hence, the harmonious unification of ends is more than merely giving “equal consideration” to the ends of all, as Hare proposed (see note 7), and it is also more than merely promoting a whole of the ends that each may set himself, as Guyer claims, for it implies actively engaging oneself in realizing each other's ends in a common effort.

Thus far, I have argued that the realm of ends has at least two meanings: the systematic union of rational lawgiving beings as objective ends, ends in themselves or persons, and the harmonious unification of all morally legitimate ends that these rational beings may adopt. In the following section, I will argue that the realm of ends has a third meaning that concerns the coherent and systematic unification of all *maxims* that the subject's moral actions are based on (i. e., in the realm of ends the maxims adopted by a subject have to be jointly universalizable). Therefore, I will first have to analyze what Kant means by saying that the formula of the realms of ends is the complete determination of the categorical imperative.

#### 4. The Unification of Jointly Universalizable Maxims

In *GMM* Kant claims there is only one unconditionally demanding moral law or categorical imperative (CI) (4:421, 436). Nevertheless, he distinguishes several (kinds of) formulas of CI: (i) the formula of a *universal law* (FUL), “act only in accordance with that maxim through which you can at the same time will that it become a universal law” (4:421; see also 4:402, 434, 437, 438; 6:389; 27:469); (ii) the formula of *humanity as end in itself* (FH), “so act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means” (4:429; see also 4:433, 436, 437, 438; 6:236, 410, 462); (iii) the formula of *autonomy* (FA), “act

---

13 Guyer, *Kant's Groundwork*, 100.

only so that the will could regard itself as at the same time giving universal law through its maxim” (4:434; see also 4:431, 432, 440, 447; 27:469); and (iv) the formula of the *realm of ends* (FRE), “act in accordance with the maxims of a member giving universal laws for a merely possible kingdom of ends” (4:439).<sup>14</sup>

In summarizing his account of CI, Kant repeats that these different “ways of representing the principle of morality” are to be seen as “only so many formulae of the very same law”, but at the same time he adds that “there is nevertheless a difference among them” (4:436). This difference concerns the fact that FRE implies a “complete determination” of CI referring to the harmonization of “all the maxims from one’s own lawgiving [to (“zu”)] a possible kingdom of ends” (4:436). As Kant argues, FRE unites both (i) the *form* of the maxims consisting in their universality, so that maxims have to be chosen as if they held as universal laws, and (ii) the *matter* or absolute end of the maxims; the latter is related to the demand that the rational being as an end in itself has to serve as the limiting condition of all merely relative ends (4:436). Obviously, form and matter refer to FUL and FH, respectively. Yet FA is also incorporated into FRE as is proven by adding “from one’s own lawgiving”. Connecting FUL, FH, and FA to one another *in* FRE prevents any merely formalistic interpretation of Kant’s CI. Moreover, FRE expresses a supplementary *communal* dimension that is not made explicit by FUL, FH, and FA separately.<sup>15</sup>

Kant connects this differentiation between form, matter, and complete determination to a “quantitative” progress from “unity” of the form of the will and “plurality” of the matter of the will to “allness” or “totality” in FRE. In the first *Critique* “allness” or “totality” is defined as “nothing other than plurality considered as a unity” (B110–1); this amounts to saying totality exists in a harmony of plurality *in* unity. FRE requires that a *plurality* of both rational beings as ends in themselves *and* of their particular—morally legitimate—ends be

14 Alternatively (4:438): “Every rational being must act as if he were by his maxims at all times a lawgiving member of the universal kingdom of ends.”

15 As Robert N. Johnson, “Kant’s Moral Philosophy”, *The Stanford Encyclopedia of Philosophy* (available at [plato.stanford.edu/archives/spr2004/entries/kant-moral](http://plato.stanford.edu/archives/spr2004/entries/kant-moral), [20 January 2009]), emphasizes: “The intuitive idea behind this formulation [FRE] is that our fundamental moral obligation is to act only on principles which could earn acceptance by a community of fully rational agents each of whom have an equal share in legislating these principles for their community” (§8).

brought into harmony under formal unifying principles of pure practical reason. Hence, FRE reveals its essential *social* or *communal* dimension: it is only in and through a realm of ends that a universally valid law can be found for a *plurality* of rational agents *united* in an ideal moral *community*.

Moreover, FRE extends the function of autonomy and dignity as *principium executionis* for executing one's moral actions to all rational subjects *as members* of a moral community. Such membership guarantees that the autonomous will can serve both as its *own* and as a *universal* law: the common laws uniting rational beings in a realm of ends have to be laws that are *equally valid for* and *equally recognized by* all members; this is possible only if all members are lawgiver of as well as subject to these laws. Further, FRE unites rational beings through moral laws regulating their reciprocal relations. This implies that *only if* one presupposes of oneself and of all others—insofar as these are affected by the actions of one another—that they do not only have their own personal ends, but also possess the capacity to and the interest for universal (self)legislation, is a *systematic* connection of *all* morally legitimate ends *possible*. Hence, it is guaranteed that all involved agents are capable of abstracting from their particular ends and of constructing a universally *valid* and universally *acceptable* maxim so that a harmonious unity of all ends can be realized.<sup>16</sup> Finally, FRE implies that a realm of ends is so regulated that, in striving for particular purposes, each member as a universal lawgiver never restricts other members from doing the same: each member has to take into account his own perspective *and* at the same time the perspective of the other members as equal self-legislators.<sup>17</sup>

Valuing FRE as complete determination of CI thus reveals fundamental insights into the truly social nature of Kant's ethics as a communal moral practice that is characterized by the unity of moral personhood understood as the union of rational lawgiving beings as ends in themselves and the harmonious unification of the morally legitimate ends that those persons may set and pursue in common. Yet the notion of moral personhood in the realm of ends has another, third meaning. Remember that FRE states “act in accordance with the *maxims* [plural!] of

---

16 As Paul Guyer, “The Possibility of the Categorical Imperative”, in Paul Guyer (ed.), *Kant's Groundwork of the Metaphysics of Morals: Critical Essays* (Lanham: Rowman & Littlefield, 1998), 238–9, notes: “... only in a community all of whose members see themselves as universal legislators and not just as pursuers of individual ends will there always be some maxim that is indeed not only acceptable to but well motivated for all.”

17 4:438. See also Allen W. Wood, *Kant* (Oxford: Blackwell, 2005), 141–2.

a member giving universal laws for a merely possible kingdom of ends” (4:439, italics added), and that FRE as “complete determination” of CI refers to the harmonization of “*all the maxims* from one’s own lawgiving [to (“zu”)] a possible kingdom of ends” (4:436, italics added). Hence, the realm of ends also entails a harmonious unification of all the subject’s *maxims*. What could this mean? In setting ends and choosing what kinds of actions are permissible to realizing those ends, a subject can adopt several maxims. FRE requires that all the subject’s maxims be jointly compatible and—even more—jointly universalizable. As Guyer makes clear, a moral agent does not merely act in accordance with one maxim considered in isolation, but rather “seeks *consistency* or *systematicity*” among *all* the maxims he adopts, and hence acts only in accordance with *all its maxims* as “comprising a coherent and systematic whole”:

a rational being will not just ask whether a particular maxim on which it proposes to act treats itself and every other rational being as ends in themselves ... and will therefore check whether *that* maxim could be universalized and if desired acted upon by everyone ...; it will also ask whether any particular maxim on which it proposes to act could be part of a coherent system of maxims on which it and all others could act.<sup>18</sup>

Only as such, does one capture the realm of ends in its fullest threefold account: the systematic union of rational beings as ends in themselves pursuing their particular, morally legitimate ends in a common effort by following the coherent system of jointly universalizable maxims they endorse.

## 5. An Ideal and Idea of Pure Practical Reason

Thus far I have discussed the threefold meaning of the realm of ends, but I have not yet analyzed its epistemic status. For Kant, the realm of ends has no empirical reality that can be proven theoretically, for we have not the least sensible experience of it. Hence, Kant calls the

---

18 Guyer, *Kant’s Groundwork*, 98. Thomas Pogge, “The Categorical Imperative”, in Paul Guyer (ed.), *Kant’s Groundwork of the Metaphysics of Morals: Critical Essays* (Lanham: Rowman & Littlefield, 1998), 204, therefore rightly concludes that FRE “demands that one extend one’s reflection to further and further such maxims while also eliminating maxims so as to consolidate the remainder into a unified system of maxims that one can will to be universally available together. In the limit, this process of reflection converges toward a complete system of jointly universalizable maxims.”

realm of ends an “ideal” (4:433, 462). Although it is our duty to realize this ideal as much as possible (i. e., to adapt the sensible world as much as possible to the intelligible world), this ideal remains always an object of mere thought for finite beings like us. While an idea can be defined as a concept of pure reason, an “ideal” concerns the “representation of an individual being as adequate to an idea” (5:232), “the idea not merely in *concreto* but in *individuo*, i. e., as an individual thing which is determinable ... through the idea alone” (A568/B596), or “the concept of an individual object that is thoroughly determined merely through the idea” (A574/B602). As such, Lequan is right in arguing that the ideal is also an intermediate between the idea and its defective sensible copies. On the one hand, an ideal shares with an idea its transcendental character, for neither the ideal, nor the idea can be represented in sensible examples—i. e., no sensible intuition can correspond to it. On the other hand, an ideal shares with objects of sensible experience that it is a concrete singular whole. Hence, the ideal is in one sense more difficult to be realized than the idea, but in another sense, more easily presented, since it is less abstract than an idea. The practical ideal is a singular thing determined by the pure rational idea. It has a practical, regulative effectiveness and functions as a perfect norm or transcendental prototype for human action. As such, the realm of ends regulates our actions by showing us the most appropriate way of moral conduct.<sup>19</sup>

Although it is a determinable individual thing, an ideal stands like an archetype that completely determines its copies. As Kant writes: “just as the idea gives the rule, so the ideal ... serves as the original image for the thoroughgoing determination of the copy” (A569/B597). In this respect, Kant mentions the ideal of the Stoic sage as an exemplar designing an image of complete conformity with the idea of wisdom. Thus, while virtue is an idea of reason generating rules for moral judgment, the Stoic sage is an ideal serving as a unique standard or model to shape our moral conduct: “we have in us no other standard for our actions than the conduct of this divine human being, with which we compare ourselves, judging ourselves and thereby improving ourselves, even though we can never reach the standard” (A569/B597). Likewise, the realm of ends can perhaps be interpreted as the only ideal of reason we can use to compare our human relations and interactions: it serves for designing a vivid image of what would be a perfect *social moral practice*, whose con-

---

19 This characterization is indebted to Mai Lequan, *La philosophie morale de Kant* (Paris: Seuil, 2001), 238.

cept is “complete in its kind, in order to assess and measure the degree and the defects of what is incomplete” (A570/B598).

Moreover, Kant also assesses the realm of ends in terms of a “practical idea” by opposing it to the realm of nature. The realm of ends concerns not “a theoretical idea for explaining what exists”, but rather “a practical idea for the sake of bringing about, in conformity with this very idea, that which does not exist but which can become real by means of our conduct” (4:436). Hence, Kant takes the realm of ends as a normative idea of practical reason to model the world (i. e., the whole of mankind’s moral interactions) as it ought to be. In *LPR* Kant reaffirms that the realm of ends is a constructive concept of morality by arguing that “a system of ends through freedom can be attained by means of the principles of morality” (28:1099), and solely by principles of morality (i. e., not by natural principles). Hence, Kant argues that, while the “realm of nature” is possible only on the basis of heteronomous, deterministic, and mechanistic laws of “externally necessitated efficient causes”, the “realm of ends” is merely possible through autonomous maxims (i. e., “rules imposed upon oneself” (4:438), and—*a fortiori*—rules imposed and generated by oneself, for otherwise these rules would not be autonomous at all). As Kant puts it in his *Lectures on Ethics* (29:629, italics added):

If I picture to myself a kingdom of natural things, ... then that is the kingdom of nature under heteronomy. But I can also picture a kingdom of purposes with autonomy, which is the kingdom of rational beings, who have a general system of ends in view. In this realm, we consider ourselves as those who obey the law, but also as those who *give* laws.

Holtman therefore nicely characterizes the realm of ends by arguing that “we must ... conceive of this community as one in which we are *governed by laws of our own making* that reflect our mutual status as ends.”<sup>20</sup>

Kant explicitly focuses on the foundational role of autonomy for the realm of ends. The autonomy that the realm of ends is based on is not to be seen solely as freedom from mere mechanical causes: moral autonomy implies not merely negative freedom as independence from (empirical-mechanical) constraints, but, on the contrary, positive freedom understood as the capability of lawgiving or self-legislation as well as the rational capacity of setting ends. Precisely by taking autonomy into account, not as a principle of mere individual self-mastery, as in the case of

---

<sup>20</sup> Holtman, 107, italics added.

the Stoic sage, but as a principle of self-legislation *within* (the framework of) a (transcendental) community of equal lawgiving members, the realization of universal unity amidst particular plurality in Kant's realm of ends is made possible: one obtains not only a union of rational lawgivers as ends in themselves, but also a non-conflicting unity of their common substantive purposes amongst a plurality of their particular relative purposes. The freedom and the autonomy involved in the realm of ends cannot be equated with absolute, lawless or unrestrained freedom and autonomy. Rather, they designate answerable or responsible freedom and autonomy that are called to account by the others' points of view, that never degrade the others' humanity as a means to an end but always value it as an end in itself, and that never can be merely individualistic or egotistic but can only have a right to exist within a (transcendental) community of equal self-legislators. In sum, Kant's realm of ends is characterized by autonomy or self-legislation, reciprocity, and mutual responsibility.<sup>21</sup>

In this essay I have argued for acknowledging the multi-layered meaning of the realm of ends as the locus for moral personhood. The realm of ends concerns the union of rational lawgivers as objective ends or ends in themselves, the harmonious unification of all morally legitimate or permissible subjective ends, and the coherence and systematicity of all the subject's maxims whereupon his or her moral actions ought to be based (i. e., the unification of jointly compatible and jointly universalisable maxims). Furthermore, I have focused on the irreducibly communal dimensions of the realm of ends. As an ideal of pure practical reason, the realm of ends is to be understood as a normative archetype for modeling an intersubjective, communal moral practice. I believe these perspectives are a valuable corrective to persistent but outdated portrayals of Kant's moral philosophy as atomistic and adverse to community life. Kant's realm of ends can therefore be seen as the locus for a communal moral practice based on the unity of moral personhood in the threefold sense discussed above.<sup>22</sup>

---

21 This point is aptly made by Korsgaard, 212.

22 I wish to thank Stephen Palmquist, Bart Vandenabeele, and the Research Foundation Flanders.



## 30. Kant's Notion of Perfectibility: A Condition of World-Citizenship

*Monique Castillo*

### 1. Perfectibility as Solidarity

Kant's cosmopolitanism has both a political and a cultural meaning. At the political level, it means the building of a free confederation of republics, of a peace alliance. At the cultural level, cosmo-political citizenship inaugurates a culture of free circulation of people, ideas, and knowledge, so that hospitality may prevail over enmity.

We speak today of globalization rather than cosmopolitanism in order to indicate that the global market creates a homogeneity of economical behaviors that can promote the unification of mankind. But the ground of globalization is technical, not ethical: it is not based on the moral will of individuals, but on the efficiency of the circulation of goods. I would like to discuss the following hypothesis: our evolutionist and scientist vision of progress has concealed the anthropological dimension of globalization; progress in the technical order of things has made us forget human perfectibility.

When we think that the progress of human civilization consists in the transition from cosmopolitanism to globalization, we often have in mind a positivistic view of progress that supposes the evolution of science and techniques can replace moral choice, that production can replace human action, that technological innovation can supplant moral creativity. But a shallow positivistic thesis leads to an absurd conclusion: it identifies progress with the desacralization of the world and the desacralization of the world with a demoralization of individuals, so that we have become less and less moral and more and more progressive.

We shall go back to Kant to understand that perfectibility signifies the original solidarity of all members of mankind (individuals, nations, and cultures). Firstly, we shall see that the perfectibility of personhood is synonymous with an anthropological solidarity of mankind, and that this is the very condition for the possibility of a future cultural

world-citizenship. Secondly, we shall see perfectibility as duty. The duty Kant expresses in the Doctrine of Virtue, to make oneself “more perfect than nature only did create you” (4:419 [§4]), means that culture claims an ethical and non-technical answer to the relationship between freedom and nature.

## 2. Perfectibility from a Cosmopolitan Point of View

The eighth thesis of the Kantian “Universal History from a Cosmopolitan Point of View” defines the political goal of cosmopolitanism as follows (8:28)

The impact of any revolution on all states on our continent, so closely knit together through commerce, will be so obvious that the other states, driven by their own danger but without any legal basis, will offer themselves arbiters and thus they will prepare the way for a distant international government for which there is no precedent in world history.

Previously, the second thesis provided its anthropological condition (8:19):

Therefore a single man would have to live excessively long in order to learn to make full use of all his natural capacities. Since nature has set only a short period for his life, she needs a perhaps endless series of generations, each of which passes its own enlightenment to its successor in order finally to bring the seeds of enlightenment to that degree of development in our race which is completely suitable to Nature's purpose. This point of time must be, at least as an ideal, the goal of man's efforts.<sup>1</sup>

In the Third Definitive Article of his *Project for the Perpetual Peace* Kant provides the ethical and juridical condition (8:358): “Hospitality means the right of a stranger not to be treated as an enemy when he arrives in the land of another.” So hospitality has both a political and a cultural meaning.

---

1 Kant, *Idea for a Universal History from a Cosmopolitan Point of View*, tr. Lewis White Beck, in Immanuel Kant, *On History* (Indianapolis: Bobbs-Merrill, 1963).

A. “Civilization” and “Moralization” in “*Cultur*”

In order to adopt a cosmopolitan point of view, we have to reconstitute the entire movement that takes into account the history of mankind before and after Enlightenment. The concept of culture has two meanings. Socialization is preceded by the first steps of culture, by the polishing of individuals: before being “civilized”, the individual is first of all “cultivated” (15:780; R1498): at this point, culture means the capacity to use tools for the preservation of life. But if we situate ourselves at the other end of the process, civilization is then overcome by *culture*, understood here in a moral sense (8:26): “The idea of morality belongs *also* to culture (*Cultur*).” Thus, the complete project of humanization extends culture beyond the civilization of Enlightenment. At the beginning, individuals surpass the state of nature and constitute society; after the progress of civilization, an overcoming of socialization toward moralization is foreseen—that is, the overcoming of the culture of skill (*Civilization*) by a culture of freedom (*Moralization*). This means overcoming the stage of the education of humanity based on discipline toward education based on freedom itself (15:608 [R1396]): “Mankind reaches his natural destination, i. e., the development of his talents, through civil constraint. We can hope he will also reach his moral destination through moral constraint.”

In its integral sense, culture includes (techno-scientific) civilization, as well as its overcoming. Here is the originality of Kant’s proposal: it is not simply a descriptive picture of successive stages, but a dynamics of the *self-overcoming of civilization by its own means*. The evils engendered by civilization create the dynamics of the need for morality and for the passage to a higher stage of development. Civilization ends up in a need of moralization, in generating the need of an ethical relation with the world. From the reign of utility we should not expect the destruction of the will, but the birth of a new need for will. A draft of the second Section of the *Conflict of Faculties* confirms the moral meaning of culture: “The prediction of a future moral success, derived from occasional reasons (partially moral and internal, partially physical and external) given to humankind, which cannot fail to intervene, proceeds also from an Idea of the practical Reason ...”.<sup>2</sup>

---

2 23:459, my translation of: “Die Vorhersagung eines künftigen moralischen Erfolgs aus den im Menschengeschlecht gegebenen theils sittlichen inneren, theils physischen äusseren, Gelegenheitsursachen (die nicht ermangeln können einzu-

From the cosmopolitan point of view about culture, what is universalizable is neither a fixed mode of civilization, determined in an authoritarian way, nor a civilization among others, but *the overcoming of itself that it carries out*. The moralization of customs and thus the capacity of willing the laws by freedom, instead of being subjected to them by force, is the collective destiny of mankind. For Kant, the civilization of Europe only shows the path; it is overcome from the inside by the universality that undergirds it.

### B. Human Solidarity

“Man cannot become man except by education” asserts Kant in his *Lectures on Pedagogy* (9:443). This is valid for each and every individual who takes part in the same collective historicity. Perfectibility means incompleteness of the individual. That means each and every work, invention, or project exists only because successors exist. Perfectibility forms a genuine link between generations. The continual progress of sciences, for example, does not grow from the death of the researchers, but from their continual birth. Perfectibility is the prolongation of a life through another, no matter if it is the life of an individual or that of a people.

Such is then the link we can establish between anthropology and Kant's moral philosophy: in each individual, humanity surpasses the person—in other words, the future of humanity crosses and exceeds the particular destiny of each person and each community. The philosophy of history, written from a cosmopolitan point of view, does nothing but highlight this *cultural solidarity* of mankind. Such solidarity is not technically producible by way of constraint, even the “soft” constraint of consumerism. Rather, its worth lies in its ethical dimension, the presence in each of us of a trans-historicity that enables us to act for an ideal community (i. e., a community that no one can possess and exhaust). In *Religion within the Bounds of Bare Reason*, Kant speaks of a “duty of mankind toward itself” (6:96).

On the moral level, a person is an end in itself. On the cultural level, the perfectibility of persons is a cultural object of respect. One passage in *Kant's Lectures on Pedagogy* asserts (9:447): “never must the education of the children be done only according to the present state, but also ac-

---

treten) geht also aus einer Idee der praktischen Vernunft in der Ordnung der Kategorie der Modalität auf folgende Art hervor.”

according to the possible better and future state of humanity, i. e., the Idea of humanity and the whole of its destination". To refer education to the future rather than to the past, or even to the present, is to act in a way that liberates individual from the domination of their own culture—a cultural “disinstrumentalism” and a way of liberating each culture itself from its ideological annexations, from its submission to possessiveness of political and religious leaders. Everyone can assert : *no one is owner of my perfectibility*, of my ability to progress, of my aptitudes for art, science, philosophy, or virtue—neither others, nor myself. The Kantian concept of perfectibility contains a specific kind of duty, a duty toward human dignity, that Kant places above human rights themselves (according to §17 of the Doctrine of Right<sup>3</sup>). There is in each individual something higher than oneself: this supreme right of humanity that can neither be exploited, nor manipulated, nor dominated, nor transformed in accordance with my free will or the free will of one another is the ultimate destination of mankind. Perfectibility is not an empirical attribute; it is only a movement, a movement whereby my educability is exceeded in the perfectibility of mankind.

To conclude the first part of this inquiry: the concept of progress is condemned today as involving a false hope—that of the total domination of nature by the human will. But the error is to consider progress in a simply technical way, like a mere series of performances, and not in an ethical way, as a meaningful project. For Kant, action consists in giving value to life, and that makes history itself understandable.

### 3. Perfectibility as Duty

According to Kant, the vocation of mankind can be expressed in this ethical exhortation from the Doctrine of Virtue: “Make yourself more perfect than nature only did create you” (6:419 [§4]). We shall now examine this duty of perfectibility according to the relation

---

3 6:270 says: “...Daher ein Mensch sein eigener Herr, aber nicht eigentümer von sich selbst (über sich nach Belieben disponieren zu können), geschweige denn von anderen Menschen sein kann, weil er der Menschheit in seiner eigenen Person verantwortlchi ist; wiewohl dieser Punkt, der zum Recht der Menschheit, nicht dem der Menschen gehört, hier <in der Rechtslehre> nicht seinen eigentlichen Platz hat.”

Kant's third *Critique*, the *Critique of Judgment*, draws up between nature and culture.

#### A. Nature acts against my goals

Section 83 of the *Critique of Judgment* shows that the natural end of mankind as a free and natural being at the same time consists in culture. The proof of this cultural destination of human nature is given by a fact of observation: nature acts against my goals; nature frustrates the goal that I am yet quite sure is the most natural end for all living beings—happiness.

When people wonder about the ultimate end of nature, they prove that nature is not an abstract being, but has the contradictory characteristics of being natural and free. When we try to understand which meaning can be given to life, we wonder about our own *raison d'être*. But we can notice that such a question would not have any sense if nature were for us a mere foreign reality, external and indifferent. In fact, it is because we are ourselves beings of nature, *beings whereby nature transforms, changes itself*, that the question of the meaning of life becomes really understandable, taking on a *human natural* meaning.

So nature acts against my aims; but, when it frustrates my goals, it also contradicts itself. It denies itself in myself when it frustrates the goals I consider as my natural purposes. This idea is rather difficult to explain, so I quote Kant (5:431):

In order to find out where in man we have to place that *ultimate purpose* of nature, we must seek out what nature can supply to prepare him for what he must do himself in order to be a final *purpose*, and we must separate it from all those *purposes* whose possibility depends upon things that one can expect only from nature. Of the latter kind is earthly happiness, by which is understood the complex of all man's *purposes* possible through nature, whether external nature or man's nature; i. e., the matter of all his earthly *purposes*, which, if he makes it his whole *purpose*, renders him incapable of positing his own existence as a final *purpose*.<sup>4</sup>

This strange relation between mankind and nature is undoubtedly something mysterious, but it is exactly what makes it possible to raise life above life, the only fact that makes it possible to give meaning to

---

4 *Kant's Critique of Judgment*, tr. J. H. Bernard, second edition (London: Macmillan, 1914); available online at [www.econlib.org](http://www.econlib.org) (The Online Library of Liberty).

life. In this context, it is not a metaphysical, but an anthropological question, a situation we experience as an intimate contradiction, a contradiction between the end and the means of life itself.

Facing this contradiction, each of us persists spontaneously (i. e., naturally) to expect from nature some gratifications we believe are due to us because we are living beings, and the most important of these gratifications is happiness. Is not happiness the most natural satisfaction of my natural reality? However, when I am reasoning in this way, I am considering nature as a merely *physical model of happiness*; I “physicalize” happiness, ranking it in the field of things, in the field of the management of things. Thus, I consider and I handle the part of nature that is in me as a mere consumer good, and I do not hesitate to request the assistance of right itself in this pursuit of happiness: if thirst claims drinking as its natural effect, does not sexuality also claim the right to have a child as a natural effect, and consequently a legitimate effect, of my desire? Do I not have a right, as everyone, to the goods of the world? Nevertheless, I must acknowledge, even if it seems very strange, that nature frustrates these natural desires, these natural aims. It forces our species to experience sterility just as easily as thirst, suffering, and death.

Faced with such sufferings, if I persist in “physicalizing” the picture of happiness, I then choose to privilege the *technical* answer that can oppose this natural frustration of my natural needs. I then will nature to produce the maximum of the effects that I expect from it, and so I become myself a mover of the process of mobilization of every exploitable energy, mobilization of all natural resources (according to the Heideggerian conception of modern technology). I then act to produce a technical submission of nature to my desires, because—let us notice—I consider nature, outside *me*, as a mere physical collection of determinisms, as a system of mechanical causes.

#### B. Culture claims an ethical and non-technical answer to the relationship between freedom and nature

But I can make a different use of my faculty to judge; I can understand that *it is inside myself* that nature frustrates my desires and that a denaturation of my technical relation to nature itself occurs within me. *It is natural for me to denature myself*, in the sense that my relation with nature is necessarily teleological and thus cultural; it has a vocation to be carried

out in a linguistic manner through symbolic systems, in cultural creations.

This forces us to maintain an ethical relation with nature and not a merely technical relation: only action, in the teleological perspective, can give value to life, as Kant says (§83, note 1). This means that in the experience of suffering, it is not a performance that is expected from us, but the capacity to give a human meaning to our behavior, a founding of meaning to our action. For example, when a woman wants to fight against sterility by recourse to the methods of assisted procreation, she gives favor to the *technical* relationship to nature; but when she chooses the adoption of an orphan, she engages an ethical relationship with nature. So, inside an individual, physical causality can be complicated by teleological causality: in this case, the origin of action is *not a cause but a reason*, an aim or an end, a plan (an intention) that separates the individual from all the purposes “whose possibility depends upon things that one can expect only from nature.”

The Kantian expression: “Make yourself more perfect than only nature did create you” means that nature cannot be considered as a reality *in itself*, as a physical thing in itself, but only as a source of values that I can give myself (i. e., the origin of reasons to act for the moral unity of the mankind as ideal horizon of the action). The final end of our species is not to subject nature to a fantastical imagination of all satisfactions, but to make ourselves “worthy of happiness” (i. e., to adopt the cultural responsibility for a possible happiness on earth). There is in me something that exceeds me, that is stronger than me: the respect due to humanity as a collective species whose destination is moral; it prohibits me from treating myself like a thing, to sell myself, to mutilate myself, to exploit myself or to destroy myself. When individuals have only technical relations with nature, when they expect the hope of a better life from nothing else than technological intervention, it shows much more *a deficit of humanism than a triumph of humanism*. If one promotes the Kantian point of view in §83 of the *Critique of Judgment*, one understands that technical performance does not delete, but generates the need *to give oneself ends* (i. e., *to create values*). Technical solutions call for the creation of ethical answers, the invention of a reason for acting, the ethical or aesthetical expression of a relationship to the world and to the others. From the ethical point of view, a technological advance does not have only the meaning of a physical performance; it requires an ethical adaptation too, an invention of symbolism in a Kantian sense, a need for an ethical anticipation of the future.



### 3. Conclusion

Creativity is required for the survival of humankind; we have to fight against cold, danger, diseases ... But, contrary to our evolutionist and scientist vision of progress, in its continuous self-invention, perfectibility includes also the resources of ethics. Without parental love and care, a baby does not become human. We are rediscovering now, with Axel Honneth,<sup>5</sup> that without social recognition no individuals can acquire self-esteem. Raymond Aron used to call “working classes” the families that give to their child a capital consisting in moral support rather than money, succeeding thus in going up on the social ladder.<sup>6</sup>

However, the term “perfectibility” grows old. We have to translate it now into a relational vocabulary. An individual or a culture can be seen today as potential relations to others rather than as isolated atoms. Perfectibility, be it physical, social, or moral, signifies that in a human being the possible surpasses the real. These potentialities would remain unrealized without being taken over by another. As we are too accustomed to understand culture in an ethnological manner (as a collection of social skills, as a conditioning), we have lost the habit of thinking of it in ethical terms (as on-going perfectibility). But, if we are to see culture as an inspiring power, we must see it as a whole consisting of not yet accomplished potentialities, potentialities that will be achieved by relations that themselves can be found with others: other cultures, other generations. Culture thus develops the need for an ethical and aesthetic link to the world and to others.

---

5 Alex Honneth, *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, tr. Joel Anderson (Cambridge: Polity Press, 1996).

6 Raymond Aron, *La Lutte des classes* (Paris: Gallimard, 1964).

## 31. Person and Character in Kant's *Anthropology from a Pragmatic Point of View*

Marc Rölli

### 1. (The Anthropological) Ambiguity of Character

With his notion of “characteristic”, or characterization, Kant introduced a deeply problematic figure of thought into philosophical anthropology that was to have far-reaching consequences. This figure of thought consists, on the one hand, in asserting the singular unity of human nature, and, on the other, insisting on empirical differences that, by definition, are differences of character, such as person, gender, nationality (*Volk*), and race. For Kant, these differences are naturally given and can be appropriately represented and treated by empirical anthropology. Kant uses the concept of character “in two senses” (7:185). The natural differences of the “*sensible* character”, he writes, pale in light of the “*intelligible* character of humanity as such” (7:229). Thus the actual and systematic meaning of the concept is not reflected in the empirical characteristics of a person, but in one’s *essence*, in a person’s “moral predisposition” (7:228). From this perspective, the empirical differences appear as systematically irrelevant determinations that can be seen as marginal in relation to the truly interesting philosophical questions. While children, women, invalids, members of non-European races, and so forth, are rightfully considered human, the claim can always (rightfully) be made that they *de facto* are not able to think and act reasonably. This is how Kantianism keeps its conscience clear—after all, one can always (in response to the widely debated accusation of “inconsistent universalism”) simply maintain the theory while adjusting the empirical claims.

In a knowledgeable and sharply argued 2007 essay, Pauline Klein-geld put forward the claim that Kant radically revised his discriminatory perspective on non-European races in the 1790s—and did so in line with his universalist moral theory. I would suggest, however, that the notion of “inconsistent universalism” is inadequate for the complexity of the issue at hand. In the following I will argue for shifting the debate

to the area of “characteristic” (characterization). Those interrogating the *consistency* of Kantian thought (inconsistent universalism versus consistent inegalitarianism) tend to presuppose that a hierarchy of the races is *logically incompatible* with the universalist assumption of the equality of human beings.<sup>1</sup> But Kant refutes precisely this incompatibility on the basis of his anthropological doctrine of characterization. It follows, against Kleingeld, that the treatment of the race question in texts such as *Perpetual Peace* that diverge from earlier writings on the subject are not at all associated with a revision of characterization as a whole. At the same time, Kleingeld’s aim of repudiating the arguments of both Kant’s critics as well as his apologists is based on the proper insight into the theoretical nexus between practical philosophy on the one hand and certain anthropological convictions on the other.<sup>2</sup> It is precisely in the context of anthropological characterization that the relationship between problematic claims about race, gender, and so forth, and universal claims about mankind can be reconstructed. The book version of *Anthropology from a Pragmatic Point of View* (1798) provides sufficient textual support for such a reconstruction.

Kant’s model for the structure of the first part of the *Anthropology* was the empirical psychology of Alexander Gottlieb Baumgarten’s *Metaphysica*. “Since there is no other book about anthropology, we will choose as our guide the metaphysical psychology of Baumgarten, a man who is very rich in material and very brief in his explanations.”<sup>3</sup> The three chapters covering the anthropological Didactic in Kant’s work correspond to the classifications of philosophy established in the *Critique of Judgment*: as empirical and pragmatic treatments of human psychological faculties they have “a priori parallels” in the three *Critiques*. The second part, the anthropological Characteristic, is different. It covers the doctrine of natural aptitude, temperament, character of the person, physiognomy, sex, race, nation, species, and, according to the subtitle of the book version, is about “cognizing the interior of

1 Kleingeld, “Kants second thoughts on race”, *Philosophical Quarterly* 57 (2007), 573–92; 574.

2 “What is overlooked by both sides is the possibility that Kant’s principles are race-neutral in their formulation, but that his racism still makes its influence felt in his theory by affecting the articulation of intermediate principles and the selection on central problems to be addressed.” Kleingeld, 584.

3 25.2:859; translations of *Immanuel Kant’s Menschenkunde oder philosophische Anthropologie. Nach handschriftlichen Vorlesungen* (1781–82), ed. F. C. Starke (1831), in 25.2:849–1203 are my own.

the human being from the exterior" (7:183). At issue here are characteristic features of human individuals or collectives, distinguished by "external", that is, natural or cultural, traits. In this sense, the second part is no longer primarily concerned with the *universal* human, our psychological powers and faculties, but with the *particular* human: the range of gender-, age-, race-, mentality-, nationality-, etc. specific identities, and thus the distinctions between people. In the following I will work out the fundamental difference, but also the internal coherence, between the concept of mankind in Kant's moral philosophy and the empirical determinations of human nature by looking at the ambivalent conception of the character of the person and the species.

## 2. Physical and Moral Character of the Person

In the first section Kant is concerned with the *character of the person*. He begins by distinguishing between two meanings that can be attributed to the word "character". Either one speaks of *physical* characters in the plural and refers to the "distinguishing mark of the human being as a sensible or natural being", or one speaks of a *moral* character in the singular, identifying a "man of principles" as a free rational being (7:185). The character of the person can accordingly be discussed in terms of *natural aptitude* (natural predisposition) and *temperament* (disposition) on the one hand, as well as *character as such* (in the narrow sense) on the other. Natural aptitude is determined by a more or less good-natured mental state (*Gemüt*) and is expressed in the feeling for "how one human being is affected by another" (7:186), while temperament is characterized by a sensible incentive to act. The latter is (also) objectively related to the faculty of desire; the former (only) to the subjective and inner feeling of pleasure or pain. Kant refers to both character forms as natural predispositions in distinction to habitual dispositions. He divides the temperaments into the traditional forms established by the ancient notion of the four humors: sanguine (light-blooded), melancholic (heavy-blooded), choleric (hot-blooded), and phlegmatic (cold-blooded). Kant makes phenomenological use of "terms of *constitution of the blood*" (7:186) when he associates feelings and inclinations "in accordance with the analogy ... with corporeal causes of movement." The physiological terms, however, are used "psychologically", not in order to identify the bodily origins of temperamental characteristics, but in order "to classify these phenomena according to observed objects" (7:187). As

*temperaments of the soul* they cannot be observed by medicine the way the constitution of the more firm body parts or the complexion of bodily fluids can be, although they “may well also have corporeal factors in the human being, as covertly contributing causes” (7:186). While Kant rejects the notion that the temperaments can be identified by analyzing the “chemical blood-mixture”, he nevertheless falls back on physiological assumptions when he divides the temperaments into the four categories. Those of feeling on the one hand and those of action on the other can thus be “connected with the excitability (*intensio*) or slackening (*remissio*) of the vital power” (7:186). This results in the opposing pairs of sanguine-melancholy and choleric-phlegmatic. As natural predispositions that correspond to designations made by race theory or “folk psychology” and that are, for instance, hereditary, the temperaments reflect a differentiated set of categories that characterize the inclination toward particular affects, passions, and feelings—including weak or pathological forms. It is in this sense that Kant speaks of the merely sensible incentives of melancholy action—of the over-ambitiousness and greed of the choleric, or of phlegmatic torpidity, as weakness and strength, respectively—as compared to the superiority of the wise man.

In distinction to the natural character, the true character is distinguished by its moral determination not to act from inclination but to comply out of duty with the moral law. Kant emphasizes that the temperament belongs to nature, as it makes something out of a person, while the *one* character is ascribed to the person as a “free-acting being” who “*makes [something] of himself*” (7:192). It follows that the character rooted in one’s way of thinking has to be acquired and cannot be given by nature. It has to be founded in a singular and exemplary act, “as it were, by an explosion” because it denotes the absolute, non-fragmentary unity of the “inner principle of conduct” (7:194). This acquisition may be an extremely difficult matter—“Perhaps there are only a few who have attempted this revolution before the age of thirty, and fewer still who have firmly established it before they are forty” (7:194)—but no less (and no more) can be demanded from a reasonable human being. The “maximum of inner worth (of human dignity)” is achievable by everyone, “possible for the most common human reason” and at the same time, “according to its dignity, ... superior to the greatest talent” (7:195).

At this point in the treatment of the actual character of the person in the *Anthropology*, reference to Kant’s moral philosophy becomes inevitable. As in the Didactic, the Characteristic also posits truly wise behav-

ior as compatible with moral action, in such a way that the pragmatic determinations are judged in light of the moral ones. This does not mean that the *Anthropology* constitutes the counterpart of moral philosophy, as Kant called for in the *Metaphysics of Morals*. That is not possible because the pragmatic anthropology is not systematically integrated into Kant's philosophy but is rather "knowledge of the world that comes after our schooling." It is a *popular philosophy* that does not claim to communicate essential metaphysical content as such (25.2:856–57): "Our anthropology can be read by everyone, even by ladies doing their toilet." Its aim, however, is to measure the empirical character against the moral one, inasmuch as the latter dictates the level and capacity of education of the former. Kant states as a positive fact that women, but also less developed races and cultures, "savages", etc., cannot become active citizens in the kingdom of ends since it is their nature to act not according to principles but according to feeling, arbitrary rules, and so forth.

The Americans have such relations in their nature that will not allow them to become perfected. Nor are the Negroes capable of further civilization, although they do have instinct and discipline, which the Americans do not. The Indians and the Chinese also seem to have come to a standstill in their perfectibility. ... We must assume that they will progress no further since they are lacking in intellect.<sup>4</sup>

Evil is inherent in human nature, that is, on the one hand, in the "animal" organism of the bodily "machine", and on the other, in psychological inertia (the inclination toward affects, passions). Both natural moments blend into each other and are specified as mixed (and unwittingly ambiguous) "natural" qualities in Kant's empirical characterization. The notion of "natural predisposition" thus turns out to be profoundly ambivalent.

The moral character on the other hand is not rightfully determined empirically, but solely by means of the autonomy of practical reason. This autonomy is presupposed in the *Anthropology* but not discussed as such. The *Anthropology* is concerned with "the qualities that follow merely from the human being's having or not having character" (7:192). Those with strength of character act according to principles (do not lie, are not hypocritical, do not break their promises, enter into the proper kinds of associations, do not pay attention to outer ap-

---

4 25.2:843, 840; cf. 838 f; my translation from Kant's "Die Vorlesung des Wintersemester 1777/78 aufgrund der Nachschrift Pillau", in 25.2:729–847.

pearances (*Schein*)), think for themselves (do not try to imitate others) and have “greatness of soul” (7:193). What is important here is that the “inner worth” of the character is beyond any “market price”; it cannot be calculated as if it were a natural quantity, does not consist of inclinations, habits, or temperaments—but can certainly have an effect on behavior or be reflected in a facial expression.

### 3. The Character of the Species

The *Anthropology from a Pragmatic Point of View* closes with the characterization of the species (*Gattung*). Kant begins by noting the impossibility of a characterization based on the natural history of the human species, as this would mean comparing “through experience” two “species of rational being”, those of an earthly and an unearthly kind; but this, as is well-known, is impossible (7:225). Thus the only thing to do is to “say that [the human being] has a character, which he himself creates, in so far as he is capable of perfecting himself according to ends that he himself adopts” (7:226). This practical faculty, however, being what actually makes the human human, remains dependent upon an undeniable empirical nature that—according to “a supreme and, to us, inscrutable wisdom” of nature—in principle corresponds to the demands of pure reason. The *animal rationale*, that is, the “animal endowed with the *capacity of reason*”, can thus become an *animal rationale*. If our humanity, what distinguishes us “markedly ... from all other living beings”, inheres primarily in the idea of turning oneself into a rational animal, the realization of this idea is understood as a process of cultivation that, accordingly, also depends on natural conditions (7:226). Human beings, in other words, must make use of their faculty of reason because they want to escape the discord of nature, and this is possible because nature made it so. We can see that the cultural existence of mankind joins together our (general) moral determination with our (particular) natural constitution. An overarching “great nature” remains at work here; we encounter it again as “universal organism” in the context of natural philosophy (originally in the work of Schelling), along with a “small nature” of all-too-human inclinations and needs that are monitored and scrutinized by reason in accordance with the level of civilization achieved.

Kant distinguishes three specifically human “predispositions” providentially instituted by nature: the technical, the pragmatic, and the moral. While the first refers to the rational organization of the human

body, the pragmatic predisposition designates the human ability “in social relations” to be civilized and educated by culture—in other words, “to come out of the crudity of mere personal force and to become a well-mannered (if not yet moral) being” (7:228). Although humans (in distinction to animals) do not achieve their vocation as individuals but only as a species (in the sense of a continuous progression toward the better), in this way the process of culture terminates in a final end founded on the moral predisposition of the human. For the moral self-awareness of the person is “already the *intelligible* character of humanity as such”; this means a person is *good* by nature (“according to his innate predispositions”), even when experience teaches us that according to one’s sensible character or “animal tendency” a person must also be considered *evil* (7:229).

The problem of the moral education of people by their equals is, in principle, unsolvable: there is no safe harbor that would leave evil behind once and for all in the maelstrom of merely imperfect life. Thus Rousseau’s notion of the unnatural deformation of human beings by culture is by no means a “recommendation to reenter [the state of nature] and return to the *woods*” (7:231, my emphasis). Rousseau, writes Kant, is merely emphasizing the “perverted” and “inappropriate” but actually existing difficulty of moving from culture to morality, instead of from “morality ... to a culture designed to be appropriate to morality”, as would alone be in accordance with reason (7:233). The guiding star of the education of humankind as a whole is nothing other than the “impotent idea of his [man’s] own reason”, posited by human beings as a regulative ideal of “Providence” in a self-transcending move (7:233); it cannot be inferred *a priori* that humankind will achieve its vocation, “but only from experience and history, with expectation as well grounded as is necessary for us not to despair of its progress toward the better.” The progress of culture is accompanied by a strengthening of a person’s moral feeling “that justice or injustice is done to him or, by him, to others” (7:229); this is why egotistical private interest becomes subjugated to the public interest, that is, the rational laws of a constitution. This leads individual consciousness to “feel ... ennobled ..., namely of belonging to a species that is suited to the vocation of the human” (7:234). Thus individuals acquire a good or actual character by following their human vocation (i. e., by making use of their freedom in accordance with their moral predisposition and fighting against their base and dishonest inclinations), although Kant holds on to the idea here, too, that the “character of a living being is [that] ... which allows its vo-



cation to be cognized in advance” (7:234). For this anthropological idea does not refer only to the moral predisposition of the species but also to the teleological organization of all “creatures”, to the inner “purposiveness” (*Zweckmäßigkeit*) of the development of the *natural predispositions*.

#### 4. The Hierarchy of Empirical Character Differences

The *concept* of character reveals a deeply ambivalent construct structuring pragmatic anthropology and constitutes a system of references that establishes a relationship between natures and cultures throughout the history of the human race by connecting psychological, medical, racial, and ethnic aspects with each other. Nature and culture are in tune with each other because they predispose persons, whose vocation brings a developmental scheme for the species into play that allows the anthropological differences of character (as both naturally *and* culturally determined forms of existence) to be arranged according to progressive stages. Despite his explicit rejection of the physiological approach in anthropology, Kant works with assumptions derived from natural philosophy, insofar as he continuously refers to the empirical character of a person’s inner nature. Individual shifts in his views on race, gender, illness and health, growth and decay, inclination and temperament, and so forth, do not threaten *per se* the conceptual construction of Kant’s characterization—nor that of the anthropology—as a whole. It may sound like a paradox, but we must ask whether the strict (formalistic) separation of the pure character in the process of delineating (universal) moral principles does not have what is at first an entirely unexpected effect within the theory: the introduction of a hierarchical structure based on the ideal of the pure faculties into the context of the “merely empirical”—a structure, furthermore, that can never be effaced by pointing to the categorical difference between the two kingdoms.

Given the situation of anthropology in the nineteenth century we must admit that Kant’s transcendental critique of the empirical science of mankind holds a comparatively rich potential. But this potential is outweighed by the enormous impact of the anthropological concepts established along with the introduction of the ambiguous notion of character. Given the history of this impact, it cannot be deemed inconsequential for Kant to separate humanity’s pure nature (or pure essence) from our empirical nature in order to relate anthropological differences to a transcendental benchmark geared to the person and to the species.

## 32. Kant and the Possibility of the Religious Citizen

*Phil Enns*

### 1. Introduction

An account of public reason that restricts the contributions some citizens can make to political deliberation appears to violate the requirements of equality and freedom, and therefore is, at the very least, problematic. The denial of some reasons within political deliberation is not only problematic for democracy, insofar as it creates an unequal burden on citizens, but for public reason itself, by restricting the freedom of individuals to contribute reasons to public discourse. Examples of such accounts of public reason can be found in the writings of Rawls and Habermas. In response, I turn to Kant for an alternative that both acknowledges the importance of public reason while respecting the fact that many religious people hold that their religious beliefs have political relevance. What we will look for in Kant, therefore, is an account of public reason that acknowledges the difference between reasoning within religion and reasoning in public, while at the same time maintaining the freedom of citizens to contribute to political deliberation the reasons they feel are relevant.

### 2. Kant on Reason and Political Association

Following Hobbes, Kant understands political life as the rational alternative to a “purposeless state of savagery” and a means for pursuing goals that could not be achieved without freedom (8:25). Reason leads people into a commonwealth with a law-governed constitution that provides citizens with the rights necessary for the freedom to pursue their own goals and purposes (8:24). However, leaving a state of nature for the freedom and security of a law-governed commonwealth does not mean that citizens have left behind conflict. Kant sees conflict as an inescapable, and even necessary, element within the development of a law-governed commonwealth. Conflict will inevitably arise between

citizens as their particular pursuits diverge and clash. The function of a law-governed constitution is to provide the means for managing this conflict so that people can be free to pursue their own ends up to the point where this pursuit imposes on the freedom of others to do likewise. Understood this way, Kant's commonwealth acknowledges in its constitution that there is no escape from the *all against all* found in the state of nature but that this conflict can be placed under external laws that provide a freedom not possible in the state of nature. Kant, therefore, identifies the ideal state as combining the greatest possible freedom under external laws possessing an irresistible force (8:22).

If reason leads people to seek out political association under a law-governed constitution in order to possess freedom to pursue individual goals, then reason requires that all people subject to those laws be considered equal before the law. Only a law that considers citizens solely as citizens can ensure that the restriction of freedom is evenly distributed thereby establishing a right. What form, then, must laws take to establish the maximum freedom and equality of all citizens as citizens? Kant answers in *Perpetual Peace* (8:350): "external and rightful equality within a state is that relationship among the citizens whereby no-one can put anyone else under a legal obligation without submitting simultaneously to a law which requires that he can himself be put under the same kind of obligation by the other person." Satisfying the criterion of mutual obligation ensures that a law creates an equal obligation for all citizens by requiring that all citizens could accept the imposition of the law on themselves. It would seem then that, according to reason, laws must have the same general form as moral duty in that what is determinant "is not the content of the will but the pure form of universal lawfulness embodied in its maxim" (8:283). Like duty, their having the form of universality rather than having any specific content or particular purpose determines the lawfulness of laws.

There appears to be, then, an important difference between Kant's position on the lawfulness of laws and those found in Rawls and Habermas, a difference that has considerable import for the consideration of religious reasons in political deliberation. According to Rawls and Habermas, the lawfulness of laws requires that both laws and their justification satisfy the criterion of universality. For this reason, both Rawls and Habermas require that at some point public reasons must be provided for political decisions. However, according to Kant, the lawfulness of laws lies solely in their having the form of duty, irrespective of content and purpose. On this account, the justification for a law is independent

of the question of its lawfulness. If the lawfulness of laws depends solely on their form, then the reasons for their being legislated cannot and must not be relevant. Following this argument, if religious reasons are given in favor of a piece of legislation, this may or may not have some bearing on the likelihood of that legislation becoming law, but it cannot have any bearing on whether that legislation has the form of lawfulness required by reason. By distinguishing between the justification of laws and the matter of their lawfulness, Kant opens the possibility that religious reasons could play a significant role in political deliberation and justifying political decisions.

The relationship between the justification and lawfulness of laws raises a more important issue regarding the very nature of political association. If, as Rawls and Habermas insist, the lawfulness of laws requires justification that satisfies the criterion of universality, then the functioning of a political sphere according to public reason presupposes citizens who are already moral and quite capable of resolving their differences. Rawls and Habermas both presuppose a condition where people pursuing incommensurable goals turn to the idea of the law as a means of producing common ground. However, on this account, it is not clear what role is played by actual laws, since the justification of all political decisions according to public reason creates the common ground needed to overcome the differences between citizens. It would appear that what is necessary for citizens to organize themselves as a democratic society is primarily the idea of the law; as an ideal of reason, this provides the means for coordinating particular political decisions. Since political outcomes are justifiable according to public reasons, they could be agreed to without actually becoming law. What then is added by such a political outcome becoming law? Furthermore, since public reason, not laws, provide the primary grounds for establishing political community, those belief systems that cannot be easily accommodated within public reason must necessarily be either exiled or reduced to second-class citizenship.

Kant rejects arguments like those given by Rawls and Habermas above, responding that one cannot organize a state by presuming the participation of moral citizens who are both able and willing to reach agreement through public reason. Instead, one must begin with the assumption of a "nation of devils" (8:366). Here we see how far, for Kant, a Hobbesian state of nature endures beyond the formation of a state. Kant argues we cannot assume that the mere fact of being a citizen is sufficient grounds for expecting individuals to act as moral political

agents. Rather, Kant argues that politics is the rational means for using the natural condition of all against all in order to establish a law-governed constitution that creates the external conditions of a moral culture. As citizens, we cannot assume that other citizens will make the rational choice to respect our freedom, but we can together, through reason, establish coercive laws that guard against behavior that would violate our freedom. In this way, we, as citizens, leave open the question of whether people are moral but only expect that they act lawfully—that is, that their behavior resembles moral behavior. Furthermore, by grounding politics in the lawfulness of laws, rather than their justification, Kant leaves open the question of how citizens justify their laws. The development and operation of a democratic state that aims to maximize the freedom of all citizens depends primarily on the lawfulness of its laws rather than on whether citizens justify these laws according to public reasons. The lawfulness of laws is independent of the justification citizens have for establishing those laws and so citizens justifying their contributions to political decisions with religious reasons does not pose a challenge to political deliberation.

Kant's distinction between the form of a law and its justification locates a further difference between Kant's account of political activity and the accounts found in Rawls and Habermas. As was noted above, Kant identifies the rational quality of a law solely in its form, so that the lawfulness of a law lies in the possibility of all citizens understanding themselves as the author. On the other hand, Rawls and Habermas argue that the rationality of a law depends on there being public reasons for accepting the law, so that it is possible for any citizen to accept themselves as the author. The difference, then, between the two arguments does not lie only in their accounts of political association but also in their accounts of reason. For Rawls and Habermas, the presence of public reasons justifying a law becomes the grounds for recognizing its lawfulness, so that the rationality of justifying political decisions is, at the same time, the rationality of the lawfulness of those decisions. However, on Kant's account, the lawfulness of laws lies solely in their form, apart from their justification, suggesting two different roles for reason. First, reason fixes the lawfulness of laws through the consideration of the possibility of their being willed by any citizen. We might call this moral deliberation in that it identifies the operation of reason in its consideration of whether political decisions can be universalized as being willed by any citizen. This is a purely formal operation of reason in that it does not take into consideration the content or purposes of the political decision. This for-

mal operation of reason stands in contrast to public reason, where the content and purposes of political decisions are taken into consideration. We might refer to public reason as epistemic deliberation, in that it identifies the operation of reason in its evaluation of the various reasons offered for justifying particular political decisions. In political deliberation, people offer different possible decisions in response to particular conditions and justify those decisions using a variety of reasons. In this political deliberation, public reason involves consideration of whether the reasons provided are good reasons. Evaluation of possible political decisions cannot rely solely on whether they could be lawful but must also consider whether they will be effective. In this way, within the political sphere epistemic deliberation, with its evaluation of justification for the substance and purposes of political decisions, is distinct from moral deliberation, with its evaluation of the lawfulness of political decisions.

If the argument for excluding religious reasons from being good reasons in political deliberation is that they undermine the lawfulness of laws, then distinguishing the lawfulness of laws from their justification removes the grounds for excluding religious reasons. However, being wrong about the role of religious reasons in the political sphere does not, of course, mean that religious reasons can be good reasons. Having distinguished between the lawfulness of laws and the epistemic deliberation of political decisions, what is needed is further consideration of public reason and the role religious reasons might play.

### 3. Kant on the Public and Private Uses of Reason

In *What is Enlightenment* Kant distinguishes between the public and private uses of reason (8:36–7). The public use of reason, he says, is general, addressing an audience that could include anyone, while the private use of reason is specific and directed to a particular outcome. In the several examples he provides, Kant describes the private use of reason as the reason that operates within the functioning of a particular office or activity and is therefore limited by the expectations and constraints of properly carrying out those duties. The government official doing their job represents the private use of reason. However, an official who offered a judgment regarding their job, either to justify its importance in the operation of government or to offer critical comments aimed at improving the effectiveness of the position, would be employ-

ing a public use of reason. This would be a public use of reason insofar as the individual is no longer operating according to the function and purpose of that particular office but addressing anyone who might listen.

It would be a mistake to think Kant's distinction between the public and private uses of reason corresponds to a distinction between similar forms of rationality. Instead, Kant's account establishes an asymmetrical relationship between the two. When the government official offers a justification or criticism of their office, this public use of reason is constrained by the private use of reason that is the logic of that office and guides the official in how they should function. The public use of reason by the official may lead to improvements in the functioning of the office, but the official remains obliged to act according to the private use of reason constituting the office, even as it adapts to the contributions coming from the public use of reason. While arguments about the work of government can be addressed to an audience that could include anyone, the arguments also address a private use of reason that guides individuals according to specific goals and purposes. Herein lies the asymmetrical relationship between the public and private uses of reason. A public use of reason is addressed to anyone but aims at contributing to the private use of reason. As Kant puts it (8:41): "Argue as much as you like and about whatever you like, but obey!" On the other hand, the improvement of private uses of reason lies in the contributions that come from the public use of reason. Only by the increased freedom that comes from removing oneself from the performance of the various duties and obligations that structure political and religious offices can reflection on the purposes and goals of political and religious activities be possible. The hope of improving religion and government, according to Kant, lies in individuals being given the freedom to make public use of their reason and, in doing so, offering better ways for the functioning of these private uses of reason (8:41). For Kant, then, the public use of reason represents an epistemic act of freedom whereby people evaluate and argue about private uses of reason in the hope of improving, not only those particular activities, but society as a whole.

Kant sees the freedom of a society being possible only when there are institutions governed by rules—that is, private uses of reason—and the public use of reason that allows for the possibility of all citizens being able to contribute to the functioning of these institutions. We cannot assume that in society other people will be rational, so there is need of rule-governed institutions, to encourage people to act rationally. The more freedom people have to contribute to the establishment of

these rules, the more likely they will freely make them their own. However, the deliberation involved in evaluating the rules of institutions, like government offices and religions, is primarily epistemic, while the rationality that occurs within government offices and religions is primarily duty-like, in that it commands obedience. Since participation in these institutions is not a moral duty, and public reason does not impose a moral duty, it is possible for people who are not government officials to employ a public use of reason regarding the functioning of government offices without being committed to acting as government officials. Similarly, it is possible for non-religious people to employ a public use of reason regarding the functioning of a religious institution without being committed to following those religious beliefs. The public use of reason requires that there is obedience present as an institutional reality, but since this use of reason is epistemic and not moral, deliberation on private uses of reason, in and of itself, does not require obedience to those institutions. The free public deliberation of private uses of reason does not demand obedience from the individual but rather aims to promote freedom in society by encouraging freedom within the functioning of institutions that require obedience. A free society encourages public deliberation on all social institutions in order to encourage people to make use of their freedom to make their own choices to obey as members of multiple social institutions.

The private and public uses of reason therefore reside on a continuum of freedom. At one end of the continuum, the restriction on freedom, given the fact that people cannot be expected to act rationally, is necessary for promoting, at the very least, rational behavior. At the other end of the continuum is the freedom to deliberate on any topic. Kant's hope is that between the two uses of reason, a balance can be found that allows for people to act and think as freely as possible but also provides enough constraints on people to force them to behave rationally.

#### 4. Conclusion

Making the distinction between public and private uses of reason in this way provides an answer to the concerns both Rawls and Habermas have regarding the role of religion in politics. Religious people can and should be fully integrated into the public sphere, feeling free to contribute any reasons they wish, all the while being able to recognize that the lawfulness of political decisions resides in these decisions being identi-



fied as the will of all citizens. In this way, the lawfulness of political decisions is no longer tied to contested reasons but rather to the rationality that structures the political realm: in short, the democratic use of private reason. The development of a democratic society, therefore, requires the increasing integration of private and public uses of reason in a way that balances the constraints of lawfulness with the freedom to evaluate any and all laws and rules. The question is, therefore, no longer whether a balance between public reason and religion is possible, but rather to what extent the two can be integrated. The greater the freedom people have to discuss and argue about religious beliefs and rituals, the greater their freedom as citizens. Similarly, the greater the freedom religious people have to discuss and argue about politics, using whatever reasons they see fit, the greater their freedom as citizens. In short, and paraphrasing Kant, a democratic society ought to encourage religious people to discuss and argue about whatever they want and however they want, all the while acknowledging the necessity of the lawfulness of laws and rules. Promoting freedom cannot be seen as an activity that can be isolated to parts of the lives of citizens. Liberal democracies that aim to promote the freedom of citizens ought to see the expansion of political deliberation to include religious reasons as expanding freedom, not only within the political sphere, but also within religious institutions.

PART VII

Persons in Ethico-Religious  
Community



### 33. Autonomy and the Unity of the Person

*Susan Shell*

From early on in his writings, Kant was concerned with the problem of self-identity and unity of personhood—not only as a theoretical conundrum but also for its moral implications (e. g., its bearing on the question of human immortality). His discovery of the principle of autonomy (in the early 1780s) altered his general understanding of this issue in important ways, issuing in theoretical perplexities that serve a series of related practical needs. That autonomy is “paradoxical” by Kantian lights is a fact not often noted. Among the many excellent Kant studies that have proliferated in recent years, few dwell upon the specific *resistance* that autonomy’s demands, in Kant’s view, necessarily provoke in us.

The *Critique of Pure Reason* had referred the question—“what ought I do?”—to a “moral” realm beyond the proper purview of critique (A805/B833). By the time he writes the *Groundlaying*, Kant has amended this view: pure *practical* reason, as he now insists, must also critically establish the fitness of its claims. Not yet prepared to present a complete “critique of pure practical reason”, Kant offers here instead a “laying of the ground” (4:392), where he identifies freedom for the first time as “autonomy” of the will. With this formula of autonomy, Kant unambiguously asserts for the first time the sufficiency of law as incentive of the will (4:433; cf. 450).

So much is relatively well known. Less understood is that the motive and aim informing the *Groundlaying* are both more modest and more ambitious than is generally recognized—more modest, because Kant does not try to derive the moral law’s validity from our practical (but not moral) presupposition of freedom, as is sometimes urged. Absent the primary orientation implicit in what Kant calls “ordinary moral understanding” (and with it, “common ethical rational cognition”), we would be indifferent to the law’s demand. In aiming to “seek out and set fast” the supreme principle of morality, then, the *Groundlaying* does not claim to give us new reason to be moral, but to release us from a moral sophistry that obscures what is “before our eyes”. It is true that Kant urges the necessity of presupposing freedom even in the case of *non*-moral action. But this is not (as is sometimes asserted)

because he means to “deduce” the moral law from a non-moral premise. Its purpose is, rather, to help overcome the dialectical resistance human reason is generically prone to have to its own law. At the same time, the *Groundlaying* is also more ambitious than is often recognized, aiming not only to establish the limits of human reason but also to indicate (from a point touching on those limits) “something” beyond.

In the *Critique of Pure Reason* the incentive of a rational will had still been dependent on some end: absent belief in the God and afterlife that make this end possible, the moral law is an object of “awe” but not of “execution.” In the *Groundlaying*, by way of contrast, the good will is determined by the law’s form alone, without reference to any (determining) end. With that new move comes recognition of a related dialectic of *practical* reason, whose burden philosophy itself cannot escape. A metaphysics of morals (*Metaphysik der Sitten*) is “indispensably necessary”, as Kant now says, not merely from “a motive of speculation”, but also because “morals themselves” remain “precarious” (*mißlich*) and “subject to all kinds of corruption” so long as their “guiding thread and highest norm of correct judgment” are lacking (4:390). The main purpose of such a moral metaphysics—the “entirely new realm/field [*Feld*]” Kant proposes to enter—is thus to secure this highest norm against forces (not yet specified) that make it “waver”.

This consideration helps explain Kant’s “method” and a related division of the work that many commentators have found puzzling. Section One (“transition from common moral rational knowledge to philosophic moral rational knowledge”) draws attention to and clarifies the pure principle contained in ordinary rational knowledge of morality. Section Two (“transition from popular moral philosophy [*Popularphilosophie*] to the metaphysics of morals”) picks up the thread with a reason that has lost its “innocence”, issuing in a “natural dialectic” that exposes it to the sophistry exemplified in *Popularphilosophie*, a powerful intellectual movement of Kant’s time. Abetted by such sophistry, worldly experience as to the force of natural inclination gives rise to doubt that the concept of duty has any object to be encountered in the world.

Kant counters that if anything can show *our* objective “assignment” (*angewiesen*) to law, it is precisely the imperviousness of its claim before the opposition of such inclination (4:425). It must accordingly not “even enter one’s mind” to derive the law’s reality (*Realität*) from the specific qualities of human nature. With this warning, moral philosophy reaches a crucial moment of self-recognition:

Here we now see philosophy placed in fact [*in der That*] in a precarious standpoint, which should be firm [*fest*] in disregard [*unerachtet*] of there being nothing in heaven or earth from which it hangs/depends or on which it is supported. Here it should show [*beweisen*] its purity [*Lauterkeit*] as autocratrix [*Selbthalterin*] of its laws, not as herald of that which an implanted sense or who knows what tutelary nature whispers/insinuates [*einflüstert*], which taken together may always be better than nothing but which never yield fundamental principles [*Grundsätze*] that reason dictates and that must have their source, and herewith at the same time their commanding authority, thoroughly *a priori* throughout: expecting nothing from inclination but rather everything from the supremacy of law and the respect owed it [*der schuldigen Achtung für dasselbe*] or, in default thereof, condemning [*verurtheilen*] human beings to self-contempt [*Selbstverachtung*] and inner abhorrence (4:425–6; emphasis added).

Like Catherine the Great, the famous “Autocratrix of all of the Russias”, philosophy must be the self-sustaining mistress of its own laws, expecting nothing from human inclination and everything from reason’s supreme authority/control (*Obergewalt*)—a setting fast of grounds that is here itself presented as an “ought” (*sollen*).

The “metaphysics of morals” that follows responds by showing that such an object is determinately conceivable in moral terms (i. e., as a member of an intelligible world made up of rational agents). Section Three (“Final step from the metaphysics of morals to the critique of pure practical reason”) uses the concept of such a world to show why proving that autonomy itself is possible (i. e., that pure reason can be practical) is a demand that an adjudicating reason can and should disregard.

If the above interpretation is correct, the argument of the *Critique of Practical Reason* is less of a departure from that of the *Groundlaying* than is commonly assumed—and certainly nothing like the “radical reversal” that Henrich and others claim.<sup>1</sup> Both works take their fundamental bearings from rational moral knowledge that is self-evident, as Kant insists, to ordinary human understanding. Neither attempts to establish the moral law’s positive validity theoretically or on the basis of some non-moral premise (such as the freedom we necessarily attribute to ourselves when we act in non-moral contexts). The clearest difference lies in their respective claims as to what ordinary moral understanding knows with

---

1 Dieter Henrich, “Die Dedeuction des Sittengesetzes: über die Gründe der Dunkelheit des letzten Abschnittes von Kant’s Grundlegung der Metpahysik der Sitten”, in Alexander Schwan (ed.), *Denken im Schatten des Nihilismus* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1975), 55–112.

unwavering certitude: in the *Groundlaying*, we know that good will, or a will determined by the moral law, is the sole object to be held worthy of esteem; in the *Critique of Practical Reason*, we know that the moral law, in commanding us categorically and unremittingly, bears immediate witness to our freedom (i. e., our capacity to obey it). The doubts on the latter score that the *Groundlaying* had principally responded to are left behind in favor of a new set of worries that bear especially on the possibility and character of rational faith. It is no longer a matter of “discovering and setting fast” the supreme principle of morality—a principle, as Kant now assures us, that “stands fast [*steht ... fest*] in itself” (5:47). The difficulty to be addressed lies, rather, in the idea of a highest good, a concept that for rational beings like ourselves is necessarily ambiguous (cf. A810/B838). Reason can arrive at an idea of the highest good either by looking backward to the ground of willing or forward to the end. Securing the proper relation between these two ways of conceiving of a highest good is the second *Critique*’s central task. For that purpose, the *Groundlaying*’s wholly negative treatment of the concept of an end of reason (i. e., the formula of humanity as an end in itself)—a treatment that served Kant’s then purpose of placing the lawgiving “ground” before us in the clearest and most unambiguous terms—will no longer do.

These differences no doubt partly reflect changes in Kant’s own circumstances, both professional and political. Kant’s new approach in the *Critique of Practical Reason* is related to these changes. His main intellectual targets are no longer amoral skeptics who abet moral “wavering” (i. e., “virtue’s enemies”, as presented in the *Groundlaying*); instead, he must take on skeptical defenders of religious faith who attack philosophy on behalf of ordinary moral life. In calling freedom the “keystone” of his entire system (5:3), Kant speaks directly to Jacobi’s charge that philosophy, consistently and systematically pursued, leads inevitably to Spinozistic fatalism.

These new objectives call, in turn, for a new description of the fundamental moral insight that is to guide practical philosophy. Kant’s appeal to the so-called “factum of reason” is a reflection of this need and of his own fuller understanding of the “ambiguity” that moral philosophy is thereby called on to address. In meeting these objectives, Pistorius’s objections to the argument of the *Groundlaying* proved a useful foil. Pistorius had taken issue with Kant for (1) the alleged obscurity of his treatment of the difference between the phenomenal and the noumenal self and (2) the alleged impossibility of determining a concept of good

will without a prior determination of a concept of the good. In responding to the former objection, Kant presents a critically rational defense of faith in immortality that replaces the failed dogmatic argument that had been advanced in Mendelssohn's *Phaedo*. In responding to the latter objection, he sets out the sole basis for belief in God that is consistent, unlike Jacobi's skeptical appeal to "faith", with the autonomy of human reason.

In both the *Groundlaying* and the *Critique of Practical Reason*, moral necessitation forces itself unbidden upon our consciousness, on the strength of our own reason and its law. Kant now makes explicit the commanding role of legislating reason. The term "factum" does not here designate a "fact" in the now ordinary meaning of something passively accepted as "given" on the basis of experiential knowledge; Kant intends it rather in the double sense implied by the Latin original (*factum = facere*) and that is retained in the English "deed". "Factum" refers both to the doing and to what is thereby accomplished. Pure practical reason "proves its reality by what it does" (5:8).

The moral law is an "act" of pure practical reason that "forces itself", as it were, upon ordinary consciousness with an "immediacy" that bears witness to its noumenal source, as borne out by the accompanying phrase *sic volo sic jubeo* ("as I will so I command") (5:31). The phrase—uttered, in its original context, by an imperious mistress to condemn a slave in utter disregard of whether he is innocent or guilty—throws the "strangeness" (5:31) of reason's factum (or "done deed") into especially sharp relief, as a deed that collapses legislation, judgment, and execution into a single moment. Reason is indeed the *Selbthalterin* of its own laws.

To be sure, *how* it is possible for law to determine the will immediately, and without regard for any (further) good intended, still remains an "insoluble problem" (5:72). The sole feeling adequate to its effect can, however, be readily described. Although the action of pure practical reason is not "felt" directly, it is negatively registered by a feeling, designated by Kant as "respect", of the sacrifice of all sensible pleasure *in totum*. Respect is at once a feeling of humiliation (or of the "infringement" of inclination and the "striking down" of self-conceit) and of uplifting self-appropriation that is literally beyond calculation (5:73, 80). Like its aesthetically sublime counterpart, respect registers supreme value indirectly (i.e., as the "Aufhebung" or cancellation of every value that is measurable). Yet, whereas the aesthetically sublime is merely "enlivening" (as Kant will later put it), respect signals the active pres-



ence of the law as what he here calls the “supreme life principle” (5:79–81, 86, 116–7).

The “reality” of freedom from a practical (i. e., moral) standpoint “warrants”, in turn, an extension of the categories to the good will conceived as *causus noumenon* (5:50, 56). Such a *causus* acts outside the boundaries of time. Its effect on consciousness is thus “immediate” (5:46, 48, 62), just as we see “without hesitation” what, morally speaking, is to be done (5:36, 30). Conscience immediately testifies to the action of the self as noumenal source. What is to be done by *us* (as phenomenal selves), however, must be achieved in time, giving rise to complications that specifically involve the (temporal) order of precedence.

A difficulty that the *Groundlaying* had not considered thus arises. If we are to act as the moral law commands, we must be able to grasp some rule-governed link between our own causality in time and the end that a good will necessarily has in view. If the object that the moral law commands is impossible “according to practical rules”, then we cannot rationally act (in time) at all, and the moral law, in directing us toward “fantastic” and “imaginary” ends, must itself be false. The this-worldly theater of *human* action (and hence pure reason “in its practical use”) puts us in the position of requiring additional assurance as to the possibility of achieving what the moral law commands, assurance not necessarily required by all rational beings as such (5:119).

Religion is thus “subjectively” necessary, not as an incentive to moral obedience, or to help effect a lively moral interest (as the *Critique of Pure Reason* and the *Groundlaying* had respectively argued or implied), but to overcome the thought—morally disabling for a moral understanding like our own—that the *aim* of pure practical reason is itself fantastic. Without the “postulates” of pure practical reason and a related faith in immortality and God’s existence, the all-important priority of ground to end is placed subjectively in question. However “immediate” the law’s determination of the will may be, human consciousness also involves a timely sense of expectation (*Erwartung*) or awaiting what is yet to come.

In sum: the paradox associated with autonomy (4:439) has been significantly recast. It is no longer specifically linked, as in the *Groundlaying*, with the good will’s indifference to whether or not the kingdom of ends is in fact to be achieved. The stated paradox is now one of “method”—that is, of how, or in what order, to proceed in determining the concept of a highest good.

Kant's earlier formulation of the paradox had undercut the argument for rational faith offered in the *Critique of Pure Reason* without replacing it. As presented in the *Groundlaying*, the paradox of autonomy reduces religion to at best a vehicle of moral anthropology, at worst an empty shell. Compelled by newly restrictive political circumstances<sup>2</sup> to meet the charge that Critical philosophy is hostile to religion, the *Critique of Practical Reason* restates that paradox in a manner that gives religion a more robust role by including rather than excluding the concern of human reason for the success of its own projects. Out of that concern, all but cast aside by the "precarious standpoint" of the *Groundlaying*, a new argument emerges for rational faith consistent with autonomy (as that presented in the first *Critique* was not [cf. A813/B841]).

But religion's newly robust Critical role is not without its difficulties. The determination of moral principles, as Kant now puts it, is a "delicate" case, for "the slightest misinterpretation [*kleinste Mißdeutung*] corrupts dispositions [*Gesinnungen*]" (5:109–10; cf. 5:31n). If reason's religious need is held to be other than "subjective", the autonomy of reason is itself denied. One can avoid such misinterpretation, it seems, only by keeping the *subjectivity* of reason's needs and the *objectivity* of reason's lawgiving simultaneously in view.

This consideration, I believe, best explains Kant's reprise of the term "moral life", appearing in various forms throughout the *Critique of Practical Reason* (5:158). Notions of moral and spiritual life had figured prominently in Kant's early lectures on ethics and on anthropology, only to be retired with his Critical discovery that the noumenal self is not theoretically accessible. Its philosophic reprise in the *Critique of Practical Reason* has the specific function of guarding against the aforesaid misinterpretation.

"Moral life" in its current application is associated with pure practical reason in its noumenal, as opposed to specifically human, state. Moral instruction accordingly occurs through cultivation of the capacity to "feel" the "lively" effect of pure moral principles at the same time that one makes due allowance for the heart's own "natural principles" of vital motion (5:155–7). The law is "foreign to the [life] element" with "which we are accustomed." Hence, the principle of moral life appears only in the "shape" of a "compulsion" to abandon that accus-

---

2 On changes in the political climate following on the ascension of Friedrich William II of Prussia, see Manfred Kuhn, *Kant: a Biography* (Cambridge: Cambridge University Press, 2001), 329–40.

tomed element for one that enables us to sustain our personhood only with “effort” and “constant apprehension of relapsing” (5:158, 151). Principles built on foundations (*Grundlage*) other than concepts produce only *seizures* (*Anwandelungen*). In sum: the moral life makes itself felt in the compulsion to ascend from a state of morally inert comfort to one enabling us to preserve our personhood only with difficulty.

The *Critique's* final evocation of “the starry heavens above us” and “the moral law within us” places both life-outlooks in sublime juxtaposition (5:161). Animal and moral life (subjective need and objective law) remain side by side and mutually exclusive. (Only with the *Critique of Judgment*, published three years later, will they merge in a morally sublime representation of human history.) The former outlook “annihilates, as it were, my importance as an animal *creature*”, endowed “we know not how” with “vital force”; for I must “give back the matter” from which (in my capacity as animal creature) “it arose.” The latter, moral outlook, raises my value to infinity by revealing the presence of a “life” independent of what must be “given back” (5:161–2).

Kant has hit upon the formula that *Religion within the Boundaries of Bare Reason* will expand on: humanity as the site where the moral life that Kant there associates with personality is able to take root, if only we are willing to make room for it. Such humanity “with or without” personality is Kant’s late version of the conundrum that, as we have seen, is coeval with autonomy as such. As Werner Hamacher observes, “Conscience [for Kant] is ... never the consciousness of the self but consciousness of the difference that splits the self of the “moral person” off from the empirical self.”<sup>3</sup> This consciousness of the difference—Kant speaks of a “doubled self” and the “two-fold personality, which is how the man who accuses and judges himself in conscience must conceive of himself” (7:438)—“can ... only appear as the consciousness that someone else has of me, and this ‘someone else’ cannot be reduced to another egological consciousness.”<sup>4</sup>

Kant’s related treatment of conscience in the *Metaphysics of Morals* raises similar concerns. After describing conscience generally (and in a way unlikely to offend traditional Christians), Kant’s treatment takes a more controversial turn (6:438):

---

3 Werner Hamacher, *Premises*, tr. Peter Fenves (Cambridge: Harvard University Press, 1996), 105.

4 Hamacher, 105.

This original intellectual and (because it is the representation of duty) moral predisposition, called *conscience*, has in itself this peculiarity: namely, that although this its business [*Geschäft*] is a business of a human being with himself, the latter, though necessitated by his reason, sees himself driven as at the behest of *another person*. For the affair [*Handel*] is here the trying of a case (*causa*) before a court. But to represent one who is accused through his conscience *as one and the same person* with the judge is an absurd way of representing a court; for the prosecution would then always lose. Accordingly, the conscience of a human being, if it is not to stand in contradiction with itself, will have to think with all duties an *other* other than itself, that is other than the human being as such, as judge of his actions. This other may be an actual person, or a merely ideal person that reason fashions for itself.

Kant's problem lies in reconciling the traditional force of conscience as a source of "fear" and "trembling" with its critically professed seat in human judgment (and related status as a wholly human and internal business). One cannot identify the sovereign lawgiver with an external "God" without succumbing (like Abraham) to heteronomy. At the same time, one cannot identify it with one's sensible self without risking heteronomy from a more obvious direction. Kant can save conscience so construed from practical self-contradiction (i. e., render it consistent with human autonomy) only by positing a "doubled self" and "two-fold personality" that remains numerically singular. The accompanying image is of a single human being who "incorporates" both *homo noumenon* and something "specifically different" (*specie diversus*). In refraining from referring to the latter as *homo phenomenon* (as in earlier works) Kant calls attention to the fact that it, too, is to be regarded as having (aspects of) personality, as any being endowed with freedom must. The unity in difference that defines a human being's two-fold personality proves to be reducible to the "specific difference" that obtains between what Kant here calls humanity's higher and lower rational faculties—that is to say, between reason as sovereign judge and a subjective self (here referred to as "our reason") that can revere the sentence without being able to comprehend fully the might that makes possible its execution (6:439n):

When the proceedings are concluded, the inner judge, *as sovereign/might-having* [*machthabende*] person, passes sentence concerning happiness or misery, as moral consequences of the deed. As to the latter quality, we cannot through our reason further pursue the inner judge's might (as ruler of the world), but are only able to revere the unconditioned *iubeo* or *veto*.

Kant's late work thus returns us to the paradox that his treatment of autonomy began with—reason as both source and subject of the law—yet with this addition: the inner judge is now the ruler of the world, albeit one whose might we cannot fully fathom. Kant's late treatment of autonomy supports moral confidence by granting morality a power of execution that is beyond human comprehension without requiring our submission to anything external to ourselves.

## 34. Religious Fictionalism in Kant's Ethics of Autonomy

*Martin Moors*

In his *Essay on Fiction and Truth*<sup>1</sup> Dieter Henrich refers to Kant saying Kant's theory of postulates, like other discourses on fiction, is one historical form that articulated the suspicion of irreality in relation to certain representations that, however, are not as such incapable of settling a form of truth.<sup>2</sup> Irreal in one respect and truth-claiming in another, fictitious representations, like postulates and works of art as well, do hold an equivocal potential. They can be used in the "as if" mode of being true. In this use these representations will effectuate the same holding-to-be-true as assured cognitions would effectuate. In their as-if-usage, though, is implied the critical consciousness of being at a certain distance from any truth-assertion whatsoever. Or, as opposed to their mere use, in full consciousness that the entire orientation of my life is concentrated in certain truth-representations, these postulated fictitious truth-representations become affirmed with the same emphasis as in genuinely ascertained cognitions. Affirmations in the modus of rational faith get a fictitious status though they are continuous with truth: the truth of what my entire moral life is unavoidably urging me to accept. The settlement of truth cannot be dismissed when it concerns the *focus imaginarius* of all objective knowledge and moral praxis, even when the imaginary status of this focus, its irreality, is acknowledged.

Some of the religious representations Kant stages in his moral philosophy can be interpreted as "good fictions". Fictitious representations are esteemed to be good either in the functional way according to their engagement as necessary *momenta* in the will's imperative moral dynam-

---

1 Dieter Henrich, *Versuch über Fiktion und Wahrheit*, in *Bewusstes Leben. Untersuchungen zum Verhältnis von Subjektivität und Metaphysik* (Stuttgart: Reclam, 1999), 139–51.

2 This essay does not explicitly examine how Henrich's interpretation critically differs from H. Vaihinger's *Philosophie des Als-Ob* (1911) or H. Scholz' *Religionsphilosophie des Als-Ob* (1921) or E. Adickes' *Kant und die Philosophie des Als-Ob* (1927, reprinted 1972).

ics, or in the postulatory way when they are treated according to a standard of being true, in their necessary relation to what is defined to be practically true *kath'exochen*. To the first class, because of their effective functional use in moral development, belong fictitious creations as a divine Obligator and a Son of God. The second class contains postulatory creations that all share in the truth of what practical reason sets out to be the highest good. Instances of this second class are the postulation of the Highest Original Good (God) and, most importantly, the definition of religion itself as a postulatory creation, according to the creative semantic power of a “seeing as”. Let us focus on three instances, one by one, and we will see how Kant proceeds, in tight argumentations, to create a figure of religious thought where a peculiar affirmation of truth is going along with the suspicion of irreality.

1. In clearest terms and at several places in his practical works,<sup>3</sup> Kant argues the necessity to think of all moral duties according to their relation to a divine will, given *a priori*. Thinking this religious momentum with regard to moral lawgiving is specified according to the distinction between being author of the law and being giver of the law. “Who propounds [*deklariert*] that a law which is in accordance with his will shall be binding [*obligiert*] on others, promulgates a law and is a lawgiver.”<sup>4</sup> The religious momentum is thought to “explain” the categorical nature of moral constraint that rests on pure reason’s universal lawgiving. This religious signification regarding the categorical modality—not the practical objectivity!—of moral lawgiving stems from an originary activity on behalf of human imagination: as Kant states in *Metaphysics of Morals*, “making moral constraint intuitive for ourselves”.<sup>5</sup> By this “making intuitive” of moral obligation, imagination’s creation points, according to Kant (6:487), “properly speaking ... to the idea we ourselves make of such a [divine] being.”

Hence, if the objective reality of moral constraint is not what became religiously signified, nor was the objective reality of the object of this so-called religious intuition deduced, nor was there, thirdly, any vacant space open within the systematic elaboration of metaphysical principles of virtue to introduce an allochtonous duty to [*erga*] God,

3 6:227, 487; 7:36–7; Immanuel Kant, *Lectures on Ethics*, tr. Louis Infield, (Gloucester, Ma.: Peter Smith, 1978), 51–2.

4 Kant, *Lectures on Ethics*, 51; see also 5:460 (§89).

5 “We cannot very well make obligation (moral constraint) intuitive for ourselves without thereby thinking of *another’s* will, namely God’s” (6:487).

then finally, as there is no objective ground to urge imagination to perform this religious hypotyposis, there merely remains, according to Kant, a *subjectively logical* ground (6:487) for this imagination's religious creation. In itself stripped of all possible practical objectivity, the created fiction is subjectively meant ("believed") to be functionally meaningful (6:487): "for the sake of strengthening the moral incentive in our own lawgiving reason."

At the service of this genuinely moral purpose, religious fictionalism becomes positively useful. However, its usage is conditioned by a reflective abstention of belief (a Kantian version of Stoic epochè) regarding its fictitious irreality. "Good fictions", as contained in religious beliefs, do not work under explicit reflective negations. Their functioning becomes morally effective only if they are acknowledged—in the "as if" modus—to share an unprethinkable divine origin. Religious fictionalism, hence, provides nothing more than a standpoint for unreflectively expecting its proper usage. Paraphrasing a statement from the Third Section of the *Groundwork* (4:458), I would say: "The concept of a divine legislator is thus only a standpoint that reason sees itself constrained to take outside the intelligible world of moral lawgiving in order to think of itself as practical for humans". The practical potential proper to religious representations, for example by virtue of the fictitious thought of a divine Obligator behind the categorical imperative for effectively "strengthening the moral incentive", does emerge from psychological dynamisms that maxims of fear and hope, obedience, satisfaction, etc., originate from. Mediated by religious empowerment, a human being's nature is thus purposively matching a moral distress (*moralisches Bedürfnis*) (19:641–2 [R8101]). Moreover and finally, what Kant must have been inspired to argue positively for the religious—be it fictitious!—transfer from moral obligation toward a divine Obligator is to my contention the following fact. By this religious transfer, the genuine moral incentive ("respect for the law") becomes intrinsically purified from maxims of self-love and, hence, moral virtuousness in general and this incentive in particular become strengthened indeed, according to their proper "rational" inner nature.<sup>6</sup>

2. A second instance of religious fictionalism in Kant's ethics of autonomy is met in Section One of Part Two of his *Religion within the Bounds of Bare Reason*, where he is picturing "the personified idea of

---

6 Cf. 19:250 (R7108): "Die Verheissung obligiert nicht, sie benimmt nur die Ausrede der Selbstliebe."



the good principle” (6:57 f). In a similar way as what he was propounding with regard to the idea of a divine Obligator, Kant now proceeds in personifying the idea of the moral prototype of humanity. In its personified hypotyposis, this idea is fictitiously made into a religious one under the name of “Son of God”. Kant presents his christology—as he explicitly also repeats with regard to his satanology—as shrouded “in a mystical cover” (6:83) intended “to make intuitive, for practical use, the concept of something unfathomable” (6:59). He transfers the hyperbolic representations of the good and evil principles into religion, where they are staged as personified prototypes, Christ and Satan, philosophically acknowledging, however, that these prototypes actually “reside in human reason” (6:63). The religious transfer of these prototypes seems merely intended for the sake of a necessary imaginary hypotyposis that, in its turn, is functionally expected to yield moral empowerment. Kant principally dismisses to link his christology with an historical account of Jesus because (6:63–4):

from a practical point of view any such presupposition is of no benefit to us, since the prototype which we see embedded in this apparition must be sought in us as well ... and its presence in the human soul is itself incomprehensible enough that we should also assume, besides its supernatural origin, its hypostatization in a particular human being.

When, on this ground, Kant keeps any historical hypostatization at a principle distance from his religious thinking of the personified idea of “Son of God”, we understand better what he propounds in his argument on “the objective reality of this idea” (6:62–6). If there is not any objective reality made available to “realize” the religious idea of “Son of God”—the “making intuitive” cannot possibly be performed by schematism between the sensible and rational—the objective reality of the christological idea is exclusively referred back to the domain (*Gebiet*) of practical reason alone. When stripped off from any specific “religious objectivity”, the christological hypotyposis of “Son of God” thus becomes a mere imaginary fictitious representation. However, that its objective reality in Kant’s argument is firmly deduced refers in fact to its (practical) truth. The christological fiction in Kant’s philosophy of religion can very well be said to be true as long as one “sees” this fiction positively functioning in the genuinely moral context of perfecting one’s moral disposition and conversion. In tune with what was stated earlier, this “seeing” in virtue of the christological fiction functioning well, should not be accompanied by the reflective consciousness of its

irreality. By reflectively affirming the fictitious character of this “good fiction” (i. e., affirming its “nothingness”), Kant’s christology becomes totally and unnecessarily disempowered from its practical/moral potency and will finally fall in the forsaken region of unmasked delusions (as Nietzsche would have viewed it). For the affirmation of its practical truth and for its efficient moral use as well, the christological fiction is totally dependent on the religiously formatting operator “seeing as”; and this, in its turn, is seeing “as if” (i. e., “as if” this religious representation would be real). Dismissing any “regular” schematism, Kant introduces a schematism of analogy (6:65) in his argument that deduces the objective reality of the christological idea *in practicis*. But neither can this analogy, nor the symbolism based on it (5:226–7 [§59]) be a means to affirm any religious objectivity regarding the christological idea. Critically restricted to its pure practical use, the *logos* of Kant’s christo-*logy* devises nothing more than a mere home-made hypotyposis. Reason thus creates for its immanent practical use a proper symbolic means functioning well for the sake of the religious operation of “seeing as” (“as if”).

How can the religious “as if”-formatting, by creating within a fictitious realm of intuition a representation of the Son of God, ever be of positive use in moral affairs? How can Kant expect (6:61) that this christological ideal of moral perfection, “which is presented to us by reason for emulation, can give us force?” Kant upholds this ideal’s moral force under two conditions. The first condition is the imaginary process of the “making intuitive” of the good principle. “We cannot”, states Kant (6:61), “think the ideal of a humanity pleasing to God ... except in the idea of a human being.” The imaginary process of producing an intuitive (christological) representative of the good principle must be a process of personification according to a mimetic rule. Kant states explicitly (6:61): “For human beings cannot form for themselves any concept of the degree and the strength of a force like that of a moral disposition except by representing it surrounded by obstacles and yet—in the midst of the greatest possible temptations—victorious.” The process of religious personification is thus explicitly represented as a “forming for themselves” in mimetic accordance *with* themselves. It is under this self-made mimetic condition that Kant thinks the christological fiction useful in moral matters. The second condition for its moral force and use fits into the specific modus of holding-to-be-true regarding the personified christological idea. The modality of this holding-to-be-true is, in this case, called practical faith. Kant states (6:62): “In the practical

faith in this Son of God ... the human being can thus hope to become pleasing to God.” This epistemic condition, faith, is the subjective self-assuring attitude of the mind whereby its represented fictitious object becomes the hold for a “self-assured trust” (6:62)—that is, the trust in “yes, we can!” Moral force is yielded by practical faith. In virtue of this faith, the moral agent “steadfastly clings to the prototype of humanity” (6:62), albeit a prototype that is both true (being “an idea which has complete reality in itself” [6:62]) and at the same time belonging to the fictitious world of religiously formatted intuitions that function well symbolically (6:65; 5:226–7).

3. Thirdly, how to interpret the semantic transformation, yielded by the operator “as” in “the recognition of all our duties as divine commands” (5:129; 6:154; 6:487), a transformation whereby Kant lets his general definition of religion generate? Using (significantly) the Latin term *instar* (instead of *erga*) (6:487), Kant denotes an equalization of ponderance between, on the one hand, the morally determinative impact on the will of the pure concept of duty, denoting “the necessity of an action from respect for law” (4:400) and, on the other, an alleged supernatural divine revelation<sup>7</sup> that would imprint on me the concept of duty. This equalization makes the semantic transformation of the “seeing as” into religion possible. The semantically tranferential move—on a basis of homology—from the rationalist toward the supernaturalist position (cf. 6:154–5), regarding the cognizance of the origin of the concept of duty, is explicitly intended by Kant to let religion put additional imperative weight on this very moral concept. The purpose for doing this is, obviously, only practical: the amplification of duty’s absolute determinative impact on a human being’s will.

Kant’s argument for this additional power put on the concept of duty on behalf of its religious trans-signification (“divine command”) is developed in the context of his inquiry concerning the possibility of the highest good, being the necessary final end of all moral endeavors. Kant states (5:129): “In this way the moral law leads through the concept of the highest good, as the object and final end of pure practical reason, to religion, that is, to the recognition of all duties as divine commands.” This argument, establishing “the formal aspect of all religion” (6:487) on the basis of the concept of the highest good, differs from the former (see example 1, above) that put forward the representation of a divine Obligator. The fictitious representation of a divine Obligator

---

7 Cf. “All authority is from God” (6:319) and 6:241.

was footed on the categorical modality of the moral imperative itself, whereas this new argumentative construction is based on moral teleology. As expressed in his Comment on Ethico-theology in the *Critique of Judgment* (5:446 [§86]):

It is at least possible—and the moral way of thinking even contains a basis for it—to form a presentation of a pure moral need for the existence of a being under which our modality gains either in fortitude or (at least according to our presentation) in range, namely, by gaining a new object to which we can apply it.

Both arguments, the argument from modality and the argument from teleology as well, construe one and the same representation of “a being apart from the world that legislates morally” (5:446). As Kant states explicitly, he does believe such religiously formatted reasoning on divine legislation of the world will yield moral strength. The religious affirmation of a divine lawgiver, ruler, and judge, is assumed to be of practical worth (5:446) “even if only for the sake of avoiding the risk of [having to] regard that [moral] striving as wholly futile in its effects and of therefore allowing it to flag.”

Regarding his own understanding of the religiously formatted postulate of God's existence, Kant makes capricious shifts from his second *Critique* via his *Religion* toward his position in the *Opus Postumum*. Related to his “realist” position in the *Critique of Practical Reason*,<sup>8</sup> Kant weakens this “realism” in the *Religion*. In an important footnote at the beginning of Part I of Section IV (“Concerning the service of God in a religion in general”), Kant defends from a practical (i. e., religious) perspective a mere “idealism” of the idea of God (6:154):

with respect to the object toward which our morally legislative reason bids us work, what is presupposed is an *assertoric* faith ... that promises a result for the final aim of religion; and this faith needs only *the idea of God* which must occur to every morally earnest (and therefore) religious pursuit of the good.

Clearly different from the former realist position, Kant now states that the *idea* of God on its own suffices to assert practically, in moral teleology, the possibility of its final end. Fashioning the concept of a final end as “a special point of reference for the unification of all ends” (6:5), both natural and ethical, for whose sake he needs the unificatory idea of God,

---

8 “These postulates ... give objective reality to the ideas of speculative reason *in general* (by means of their reference to what is practical)” (5:132).

Kant actually is construing a moral parallel to his theological thinking in the first *Critique*. In his speculative theological thinking, the Ideal of pure reason (God) is thought to be regulative, at the service of an “as if” transcendental fiction with regard to (objects of) experience constituting an absolute unity (A672–3/B700–1). In such a functioning of the ideas, they direct all the understanding’s rules to what Kant significantly calls a *focus imaginarius* (A644/B672) or heuristic fiction (A771/B799). No objective reality whatsoever is needed *simpliciter* with regard to the theological idea of reason, neither in its speculative nor in its practical or religious respect, in order to let this imaginary construction function well *secundum quid*. According to Kant, the fiction functions well “only as an unique standpoint” (A681/B709) that alone one can adopt, as through a schema (A682/B710), to represent the sum total of all duties converging into one highest good.

In his *Opus Postumum* Kant construes “a totally new Idealism”<sup>9</sup> with regard to the role of the transcendental ideal, being *ens rationis ratiocinatae* (A681/B709). Contrary to figments of the brain (*Hirngespinnste*), these *entia*, created by reason, do function rationally. Consequently, Kant confers on them, in their objective relational functioning with regard to what either can be cognized or ought to be morally realized, transcendental truth. In this sense Kant, in his *Opus Postumum*, sees no contradiction between “being created by reason” (*Gedankending*) (cf. belonging to the table of *Nothingness* [A292/B348]) and assuming its existence (merely *in intellectu*, not *in re*). As V. Mathieu points out, God’s existence is assumed as “the transcendental condition of unity of ‘practical’ experience, hence of a unity *secundum quid* (according to duty), not *simpliciter*.”<sup>10</sup>

Assembling fragmentary findings from the first fascicle on the theme “There is a God”, we can catch a glimpse of Kant’s latest religious fictionalism (in its transcendental form, with respect to its truth) in statements like these:

The existence of such a [divine] Being can only be postulated in practical respect, namely, [in] the necessity to act as if I would be subjected to this ... guidance recognizing all my duties as divine commands (*tamquam non ceu*).

---

9 Vittorio Mathieu, *Kants Opus Postumum*, ed. Gerd Held (Frankfurt a. Main: Vittorio Klostermann, 1989), 211.

10 Mathieu, 258.

Hence, in this formula is not postulated the existence of such a Being, which would be contradictory.<sup>11</sup>

[God] is not a hypothetical concept sustaining other propositions but it is thought as being for itself (absolute) [*per se*] present [*bestehend*] though by this it is not stated as if such a Being would exist.<sup>12</sup>

### Retrospection and interpretation

From “within the bounds of bare reason” Kant stages, as we saw, three religiously defined entities: a divine Obligator, the Son of God, and God as a beneficent ruler of the world. The semantic class these “objects” belong to, or the text these figures plot, explicitly suggests a peculiar kind of reference. As antonyms to what is real or factual or historical, these objects, functioning on the religious stage, can be interpreted within a referential theory of modal fiction.<sup>13</sup> Though the *modi* of being or ontological status of these “objects” is explicitly withdrawn from every possible real or objective referential *correlatum* in an actual world, Kant’s religious discourse on these objects is nevertheless truth-functional. The truth-functional performativity of Kant’s discourse on religious objects—objects within a textual alternative possible world—does not work on behalf of *logos apophanticos*. In tune with Aristotle’s celebrated statement on prayer—prayer “is a sentence but is neither true or false”<sup>14</sup>—Kant as well in his religious discourse has blocked this entrance to predicative propositional truth. However, his discourse on (fictitious objects of) religion remains open to truth, according to Ryan’s “principle of minimal departure”.<sup>15</sup> This means, in Kant’s case, that those religious objects within a textual alternative possible

11 22:116, quoted in Mathieu, 261.

12 21:36, quoted in Mathieu, 263.

13 See M. L. Ryan, *Possible Worlds, Artificial Intelligence, and Narrative Theory* (Bloomington: Indiana University Press, 1991), 13–61.

14 *De Interpretatione* 17a4–5, tr. Jonathan Barnes in *The Complete Works of Aristotle*, vol. I, Bollingen Series LXXI.2 (Princeton: Princeton University Press, 1985), 26.

15 “This law states that we reconstrue the central world of a textual universe in the same way we reconstrue the alternate possible worlds of nonfactual statements: as conforming as far as possible to our presentation of AW [the actual world]. We will project upon these worlds everything we know about reality, and we will make only the adjustments dictated by the text.” (Ryan, *Possible Worlds, Artificial Intelligence, and Narrative Theory*, 51.)

world must minimally be linked to the factual world of moral practice as if this world would be dominated by “real” religious objects of dogmatic faith. The principle of minimal departure thus prevents that “any attempt at interpreting fiction would turn into a ludicrous activity.”<sup>16</sup> But does this principle, in Kant’s case, not lead us to presuppose a theological dogmatics *outside* the bounds of bare reason? Kant himself was confronted with this question in his discussion with Jacobi. In a letter to Jacobi (August 30, 1789),<sup>17</sup> Kant does not formally negate such a presupposition with regard to his religious discourse. Yet, he declares it to be an issue offstage [*Nebensache*] to his philosophical investigation. Kant states (11:74): “Whether reason, in order to arrive at this concept of theism, could have been stirred only by something taught by history, or perhaps by some ungraspable supernatural inner influence, is a question merely offstage. It does, namely, concern only the origination and arising of this idea.” After having put offstage all historical and dogmatic faith with regard to religious representations, Kant treats them as (fictitious) objects of theoretical reason and inquires their correctness and validity merely by practical reason.

These reflections do bring us to a last short remark concerning the problem of duplication or simulation. As far as one would take into account the offstage religious dogmatics as having been the historical origin of the arising in the mind of these religious representations, one could very well say that Kant in his *Religion* is counterfeiting (or duplicating) the historical/dogmatic originals. On the other hand, examined onstage within the bounds of bare philosophical reason, the same religious representations can be interpreted as simulations by a Kantian artificial intelligence. From within his autonomy-based doctrine of duties, he is creating a verisimilitude of religious truths as mental constructs in a textually alternative possible world.

---

16 Ryan, 14.

17 See also his letter to Lavater (April 28, 1775) (10:168–70).

## 35. Respect for Persons as Respect for the Moral Law: Nicolai Hartmann's Reinterpretation of Kant

*Predrag Cicovacki*

### 1. Animality, Humanity, and Personality

In this essay I am interested in examining Kant's conception of person and personality in terms of respect for the moral law. This is not the only way to understand Kant's conception, but it may well be the most important one. My presentation will consist of two parts. I will first give a brief clarification of the relevant concepts of person and personality. Then, in the light of some remarks by Nicolai Hartmann, I will examine the notion of respect, as well as related issues with regard to what it means to show respect to others. This part will raise more questions than it will offer answers. I hope these questions will lead to further considerations of Kant's rich, yet never fully developed conception of person and personality.

Kant's consideration of person and personality falls within the framework of modern philosophy, yet it also has its own peculiar characteristics. Following a broadly Christian tradition, Kant accepts the fundamental dualism in human nature and associates the seat of personality with the soul, not with the body. Following Descartes, the focus narrows down from the soul to the mind, and many modern philosophers base ethics on the philosophy of mind. In accordance with the modern preoccupation with the mind, in Kant's philosophy personality can be understood in connection with the transcendental unity of apperception, or the unity of all theoretical and/or practical activity, or, perhaps most importantly of all, in terms of autonomy.

Iris Murdoch remarks that, according to the Cartesian tradition, an agent is pictured as an isolated principle of the will. She adds that in this tradition, "the agent's freedom, indeed his moral quality, resides in his choices, and yet we are not told what prepares him for the choices."<sup>1</sup>

---

1 Iris Murdoch, *The Sovereignty of Good* (New York: Routledge, 1970), 52.



Kant's theory of personality does not seem to suffer from this omission. In *Religion within the Bounds of Bare Reason*, Kant introduces the concept of personality in the context of his discussion of several predispositions. The term "predisposition" (*Anlage*) is Kant's way of talking about basic human nature as it is prior to any actual exercise of freedom. The three original predispositions are those of animality, humanity, and personality. These predispositions correspond, roughly to: (1) physical love that provides for our preservation of the species; (2) self-love that is both physical and rational, producing the inclination to "acquire worth in the opinion of others"; and (3) "the capacity for respect for the moral law" as a sufficient incentive for the will (6:27).<sup>2</sup>

Let us now consider these three dispositions in terms of what deserves our proper respect. Kant does not have much admiration for animality. He claims, in no unclear terms, that "Life as such ... has no intrinsic value at all ... it has value only as regards the use to which we put it, the ends to which we direct it."<sup>3</sup> Is it, then, humanity that deserves our respect? Or should such respect be reserved for personality only? I do not believe this issue is clearly resolved in Kant's philosophy. What is clear is that, if human life is to gain value, or be treated as valuable, this must happen at a higher level than that of animality. Yet, at least on the surface, the next higher level, that of humanity, is the most controversial of the three predispositions. This is also where Kant's interpreters differ the most. While Christine Korsgaard, for instance, equates this predisposition with "a capacity to set ends", Yermiyahu Yovel understands it more broadly in terms of "finite rational being" that for him also includes the relevant "unsociable sociability".<sup>4</sup> Whether understood narrowly or broadly, it is clear that for Kant humanity is a precondition for personality, or the state of morality. Humanity in itself is not necessarily an actual moral state, but is at least required for its possibility.

---

2 Kant, *Religion within the Limits of Reason Alone*, tr. T. H. Greene and H. H. Hudson (LaSalle: Open Court, 1960), 22–3. My presentation here follows Gordon E. Michalson, *Fallen Freedom* (New York: Cambridge University Press, 2008), 38.

3 Kant, *Anthropology from a Pragmatic Point of View*, tr. Mary J. Gregor (The Hague: Nijhoff, 1974), §66.

4 See Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Harvard University Press, 1996), ch.4, 106–32, and Yirmiyahu Yovel, "Kant's Practical Reason as Will: Interest, Recognition, Judgment, and Choice," *The Review of Metaphysics* 52 (1998), 274–94.

It seems that this possibility is sufficient to assign to human beings an absolute value—the value of dignity and autonomy. In various contexts, Kant formulates his ideas regarding the value of human beings (as moral beings) in different ways. Most of them seem to converge in one point: respect for the moral law. In Lewis White Beck's formulation:

Personality ... is an Idea of reason, and personality is not a given. We are persons, but no finite sensuous being is fully adequate to the Idea of personality. In human nature, considered empirically, we find at most only a "pre-disposition to personality", which is the capacity for respecting the moral law and making it sufficient incentive for the will.<sup>5</sup>

## 2. Personality and Respect

I interpret the claim that personality is "an Idea of reason" in the following way. Morality conceives of a world that does not yet exist in nature and seeks to actualize it by acting in the given world according to the laws of the possible one. In Kant's words, "[moral] law gives to the sensible world, as sensuous nature (as this concerns rational beings), the form of an intelligible world, i. e., the form of supersensuous nature, without interfering with the mechanism of the former" (5:43).<sup>6</sup>

The key to understanding Kant's ethics is in the relationship of reason and will, and this relationship is marked by the distinctive feeling of respect (*Achtung*). Respect is "a positive feeling not of empirical origin ... which can be known a priori ... a feeling produced by an intellectual cause." Put differently (5:78): "Sensuous feeling ... is the condition of the particular feeling we call respect, but the cause that determines this feeling lies in the pure practical reason." It is clear from Kant's remarks that this respect can have no other than a moral ground. Furthermore, respect applies to persons only: "All respect for a person is properly only respect for the law ... of which the person provides an example" (4:401).<sup>7</sup> What is less clear is exactly how respect applies to any person: how can we show such respect?

---

5 Lewis White Beck, *Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), 227.

6 Kant, *Critique of Practical Reason*, tr. L. W. Beck (Indianapolis: Bobbs-Merrill, 1956), 44.

7 Kant, *Grounding of the Metaphysics of Morals*, tr. J. Ellington (Indianapolis: Hackett, 1974), 14n.

This question deals not only with the ground for respect, but with its scope, as well as with its negative and positive manifestations. If respect is something other persons do not have to earn, or cannot fail to earn (simply because they are rational beings and moral agents), does that turn respect into a moral *obligation*? Must other persons *always* be respected? Must *all* persons always be respected? Must all persons be treated with *equal* respect, or is there a difference in the *degree* of respect? Must each person, in order to be respected, be treated in the *same way*? How about the *individual differences* between persons? Or their actual behavior? Are they totally irrelevant? Does a rapist deserve the same respect as a morally virtuous person? Does a murderer?

Many ethicists, some of whom are Kant's scholars, do not believe the equality of all persons entails that each of them must be treated the same as every other person. Treating people with respect, they argue, does not require treating them equally. I believe this approach is plausible: there is no reason to think of Kant's view as being so rigid, or as providing some kind of unchangeable, *a priori* calculus. Yet if the difference in treating people with respect is possible and allowed, what accounts for the difference and how is the difference to be respected?

Perhaps a plausible way of approaching this latter question is by making a distinction between the negative and the positive aspects of respect. The negative aspect of respect should be invariable, and in any minimal sense of respect should involve refraining from regarding or treating other persons in morally inappropriate ways. For example, in accordance with the second formulation of the categorical imperative, other persons should never be treated as means only. This should hold true, regardless of their behavior, or motives for their behavior. This should hold true because every other person must be treated as an end in itself, as a free rational being possessing dignity.

The positive aspect of respect, however, must take into account individual differences—not all individual differences, but the relevant ones. The relevant differences must involve a person's actual behavior and motives for such a behavior. A rapist and a murder who intentionally harms others must be treated differently—although with respect—than a morally virtuous person. A morally virtuous person is shown respect in a positive way, because such a person, as a free and rational moral agent, acts from the conception of duty; a rapist or a murder does not. By punishing (say) a rapist we obviously treat such a person differently than the one who is morally virtuous, yet by punishing

such a person we also treat him or her with respect insofar we assume that, as a rational and free agent, this person could (and should) have chosen to act differently.

It is hard not to wonder, however, if those are all the differences we should take into account. We often associate personality with individuality—besides sharing a lot in common, we all have individual differences and those differences are important for whom we are. Could not Kant's ethics take such differences into consideration? Should it not?

### 3. Personality and Individuality

Nicolai Hartmann has pursued these questions more persistently than anyone else. He has also come up with some ingenious remarks worth mentioning here. Hartmann relates Kant's treatment of persons in terms of respect for the moral law with the first formulation of the categorical imperative. He points out that, insofar as this formulation affirms that the moral test for every action is whether its maxim could at the same time be a universal law,

there is evidently something here which in principle man as a personality cannot will. Rather must he at the same time will that over and above all universal applicability there should be in his conduct something of his own, which no other in his position ought to do or need do. If he neglects this, he is a mere numeral in the crowd and could be replaced by anyone else; his personal existence is futile and meaningless.<sup>8</sup>

Hartmann believes such a concern can even be formulated as a law that would run contrary to Kant's formulation of the categorical imperative. This law would say: "So act, that the maxim of thy will could never become the principle of a universal legislation without a reminder." One might also express it in this way: "Never act merely according to a system of universal values but always at the same time in accordance with the individual values of thine own personal nature."<sup>9</sup>

There is clearly a conflict here. As Hartmann would put it, following Kant, there arises an antinomy that cannot be fully removed. Perhaps the most surprising aspect of Hartmann's treatment of this whole issue is that he does not believe this antinomy creates any essential prob-

---

8 Nicolai Hartmann, *Ethics*, vol.II, tr. S. Coit (New York: Macmillan, 1932), 357.

9 Hartmann, 357.

lem for Kant's account of persons and respect for the moral law. On the contrary, it complements Kant's account and makes it richer. Here is how Hartmann comes to this conclusion. The real demand of Kant's categorical imperative is:

I ought so to will, as under literally the same circumstances everyone else ought to will. But "literally the same circumstances" includes the peculiar nature of my individual ethos. The imperative, accordingly, when the complete structure of the case is born in mind, not only excludes the moral justification of a will exactly the same in others, but it positively demands also the unique factor in my own will, without prejudice to the classification which brings my will and that of others under a rude uniformity of the Ought. The Ought allows unlimited scope for an individually articulated will.<sup>10</sup>

Hartmann concludes from this that individuality could not and should not be excluded from moral considerations, nor could or should it be excluded from the proper understanding of what it means to be a person and show respect for other persons. Although this may be an unintended consequence of Kant's theory, or perhaps simply something overlooked by Kant, Hartmann concludes that "the categorical imperative has within itself its own opposite. It involves its own converse. Its limitation lies not outside of it, but in it."<sup>11</sup>

#### 4. Personality and Love

While we are dealing with Hartmann's revisionist reading of Kant's theory of personality and respect for persons, a few more points may be mentioned. As is well known, Kant strictly opposes acting from pity or love. Such acting would effectively eliminate our respect for the persons we pity or love as free rational agents. Emotions, such as pity or love, should not be taken into account—they would precisely make it impossible to treat other persons with the (negative) respect they are entitled to and that they as moral agents cannot fail to earn. Hartmann argues that Kant thereby unnecessarily restricts his theory and somewhat distorts both the nature of personality and respect for other persons. Here is, briefly, what Hartmann has in mind.

---

<sup>10</sup> Hartmann, 359.

<sup>11</sup> Hartmann, 360.

Kant's moral philosophy is often interpreted in terms of his opposition to "inclinations" and his focus on categories such as justice. There may be good reasons for that. If, for example, we think of justice in comparison to love (e.g., brotherly love), there can be an antinomical opposition between them. In Hartmann's succinct formulation, "Justice may be unloving, brotherly love quite unjust."<sup>12</sup> Following the Stoics, Kant treats love as a pathological inclination. Yet this is not the most important aspect of love, nor its ethically relevant conception. Instead, love can be understood in terms of one's disposition or intention. Taken in this sense, love is not only close to what Kant calls good will, but it is also essentially relevant for our treatment of other persons and our respect for them. As Hartmann puts it: justice "joins person to person, but only surface with surface. ... Brotherly love binds far more deeply ... Personal love, however, unites forthwith innermost depth to innermost depth, overleaping the surfaces."<sup>13</sup>

## 5. Looking beyond Kant's Conception of Personality

Hartmann's remarks should lead us to reconsider Kant's understanding of respect for persons. Kant distinguishes three predispositions—of animality, of humanity, and of personality—and seems to be willing to treat respect for persons primarily in terms of the disposition of personality. This is so because of the ties between this disposition and the moral law. As quoted earlier, Kant maintains: "All respect for a person is properly only respect for the law ... of which the person provides an example." Following Hartmann's remarks, my contention is that such a conception is too narrow. If, because of its connections with the moral law, the disposition to personality is the only proper object of respect, too much is left out, even in the moral context. As Hartmann points out, love cuts deeper than justice and other values that would be regulated by the moral law.

There may, then, be two options open. One of them is significantly to broaden the range of respect, certainly to include the predisposition to humanity, maybe even to include some aspects of animality. The second option is to introduce two different kinds of respect—one in a narrow moral sense, focusing on personality and the moral law, and one

---

12 Hartmann, 271.

13 Hartmann, 377.

broader, including the whole range of human predispositions and abilities. I am suspicious of the ground that such a line could be drawn on. If personality is only “an idea of Reason”, it seems to be too constricted to stake the most important form of respect on something of that kind. As Iris Murdoch pointed out, “the agent’s freedom ... resides in his choices, and yet we are not told what prepares him for the choices.” The moral law deals with freedom and the agent’s choices. To understand what prepares such choices, Hartmann offers a plausible suggestion, that we need to take into account all of the agent’s rational and non-rational dispositions. Is it, then, not the entire human being that deserves our respect?

## 36. The Unity of Human Personhood and the Problem of Evil

*Aleksander Bobko*

The following thesis, in my opinion, appropriately expresses the general message of Kant's Critical philosophy: rational cognition is not only the passive discovery of the *logos* within the surrounding reality, but also—perhaps mainly—the process of gradually ordering the elements making up the diversity of the universe. Thinking is thus a power that allows reason to shape the outside reality according to its own rational criteria. That is why human beings—the real agents of such reasonable activity—can be regarded as persons: the unity of human personhood is based and rooted in rationality. Let us consider this thesis more closely. We will focus on practical philosophy only, interpreting Kant's ethics as a specific “philosophy of goodness”.

It can be stated without much risk that the most common moral intuitions are expressed in the following two basic convictions. First, the greatest good a human being can achieve in life is happiness, and the pursuit of happiness is the main, if not the only, motivation behind human actions. Second, this pursuit of happiness is realized in the effectuation of particular goals that are beneficial for the human being. Thus, the degree of fulfillment of intended goals seems to constitute the most obvious and objective criterion for evaluating the quality of human actions. In short, morally justified behavior is useful behavior.

This common intuition found its expression in most traditional systems of ethics, especially that of Aristotle. The Stagirite wrote in the opening sentence of his classic work, *Nicomachean Ethics*, that “every action and pursuit is thought to aim at some good; and for this reason the good has rightly been declared to be that at which all things aim.”<sup>1</sup> Later on in his considerations he arrived at the conviction that happiness is a human being's highest goal, “for this we choose always for self and

---

1 Aristotle, *Nicomachean Ethics* (e-text version based on the W. D. Ross translation), 1094a/1.



never for the sake of something else.”<sup>2</sup> A continued development of this basic intuition is to be found, with the awareness of some important differences, in traditional Christian ethics, as well as in many other modern systems.

Kant questions the self-evident nature of moral intuition understood in this way. To be sure, goodness remains a central concept of his ethics—as we will show, its importance grows even higher there—but its character changes fundamentally. He explicitly identifies goodness with the concept of good will. He writes of this in the famous opening sentence of *Fundamental Principles of the Metaphysic of Morals* (4:393): “nothing can possibly be conceived in the world, or even out of it, which can be called good, without qualification, except a good will.”<sup>3</sup> Even a superficial comparison of this sentence and the opening statement of *Nicomachean Ethics* reveals deep differences between the two philosophers’ ways of thinking. For Aristotle, good “exists”—the world (being) is good—and goodness is, in a sense, an object unto itself, attracting the attention of human beings. In this sense all human wishes, thoughts, and desires are directed at the good as a goal to be achieved. For Kant, goodness does not yet “exist”; reality as such is neither good nor bad. Human action is what gives value to the world: goodness may appear in the world as the result of appropriate human activity, or it may never appear, if such activity never takes place. Thus, the essence of goodness is to be found inside the human being: goodness at its source is “good will”, an appropriately formed structure of the human person as a rational being.

In comparing further these key texts of Western ethics we come across another difference. Kant goes against the common belief, shared to some extent by Aristotle, and to an even greater extent by the philosophical tradition he initiated, that the moral quality of the will is to be measured by its effectiveness in achieving goals (4:394): “A good will is good not because of what it performs or effects, not by its aptness for the attainment of some proposed end, but simply by virtue of the volition; that is, it is good in itself.” This statement clearly undermines Aristotle’s thesis that the good is the aim of all activity; Kant proposes a completely different basis for ethics.

---

2 Aristotle, 1097b/1.

3 Immanuel Kant, *Fundamental Principles of the Metaphysic of Morals* (e-text version based on the Thomas Kingsmill Abbott translation).

First of all, Kant's ethical theory marginalizes the significance of the effects of human actions. The moral quality of an act is not determined by whether or not it allowed for the intended goal to be effectively achieved. The deciding factor is simply "volition"—that is, the principle that determines the will to act. To use more modern language, what decides whether an action is morally justifiable is the quality of the motivation, not the achieved results. The most serious consequence of such a turning away from the outside world to focus on internal motivation is that happiness is no longer to be understood as the final aim of human activity. In Kant's ethics the traditional question of "How should we live in order to achieve happiness?" loses its significance. Practical philosophy is not concerned, at least not directly, with questions of how to benefit mankind or society, but with something much more important—as we shall see, this is the realization of goodness, identified with rational order, with rationality as such. Paraphrasing the basic question of classical ethics, Kant formulates his own question in the following way: How should we live in order to be "worthy of happiness"?

The premise justifying this position is the utterly serious treatment of the fact that human beings have the power of reason at our disposal. According to Kant, it would simply be preposterous if such a perfect instrument as reason served such an ultimately banal end as happiness (4:395): "If the proper object of nature were its conservation, its welfare, in a word, its happiness, then nature would have hit upon a very bad arrangement in selecting the reason of the creature to carry out this purpose." The ability to use reason is something that ennobles mankind greatly: it makes human beings capable not only of actions motivated by their inclinations and needs, but of disinterested behavior as well. For the effective fulfillment of the intended goals, instinct would suffice; reason is much too subtle and complicated a mechanism. The strongest expression of this "depth of reason" is the generation of an internal sense of duty, self-evident to any rational being, that necessarily obliges us to do what is right, regardless of the benefit to oneself. The product of reason's activity is supposed to be goodness that takes, at its source, the form of good will (4:396): "Its [reason's] true destination must be to produce a will, not merely good as a means to something else, but good in itself."

Our attention must now focus on the interaction between reason and will—that is, on the laws whereby reason determines the will to act. The concept of law plays a particularly significant role in Kant's philosophy. He understands nature, or the universe in the broadest sense of

the word, to be “the existence of things under laws” (5:43).<sup>4</sup> In this sense, all that exists is connected with some law. Yet within this law-governed mechanism mankind holds a unique place, as Kant says in *Fundamental Principles of the Metaphysic of Morals* (4:412): “everything in nature works according to laws. Rational beings alone have the faculty of acting according to the conception of laws, that is according to principles, i. e., have a will.” Mankind is not governed by the law automatically, but in a sense, mediates this law: human beings have a will, the faculty of somehow evaluating the law that is presented by reason.

What is the nature of the law appearing in the “space” between reason and will, and by what means does reason actually influence the will? At the very beginning of the *Critique of Practical Reason* Kant introduces this problem in the following way (5:19): “supposing that pure reason contains in itself a practical motive, that is, one adequate to determine the will, then there are practical laws; otherwise all practical principles will be mere maxims.” The activity of the will can be based on two significantly different schemes: the means of determining the will to act can take the form of either a law or a maxim. What does this mean precisely?

We have already said that the will is a faculty with a certain creative power; its actions can bring specific results into actual existence. The will is influenced by many factors, trying to give direction to its energy. Perhaps the most obvious determinants, at least when it comes to the will of a person as a “finite rational being”, are inclinations or desires connected with the senses. They determine the will to act by presenting a specific goal that needs to be achieved. In the case of sensual determination, the goal, as the effect of action, most commonly fulfills a specific human need. The imperative addressed to the will is only hypothetical here, limited to this single case. It says: if you want to achieve the desired result, use these specific means. If we wanted to subsume all the imperatives of this type under one formula, we would have to say that they all hold happiness as the prime aim of all human activity. There is just one problem: happiness cannot be defined precisely, so these hypothetical imperatives can only be analytical-practical judgments that do not pretend to the status of necessary judgments. If we agreed that the will is determined in this way, then practical principles

---

4 Immanuel Kant, *Critique of Practical Reason* (e-text version based on the Thomas Kingsmill Abbott translation).

would only be maxims. Maxims are principles applied subjectively, only to certain isolated cases; they cannot be said to have an objective, universal relevance. Ethics based on maxims would be a collection of “words of wisdom” that would suggest—on the basis of experience, both individual and collective, passed down from generation to generation through tradition—the most pragmatic solutions, effectively leading to the desired goal.

Kant, however, is primarily interested in ethics based not on maxims but on necessary laws. Yet these are possible only on the condition that reason determines the will directly, without any mediating elements. This would mean reason contains in itself the basis of this determination, and so does not have to present the will with any ends to be achieved through its actions. This ordering by reason also takes the form of an imperative—the will has to be forced to act—but this imperative is categorical rather than hypothetical. Its specificity lies in the fact that, disregarding all possible effects, it defines only the means of determining the will—that is, the form of the principle whereby the will shapes its volition. There is only one such imperative according to Kant (5:30): “Act so that the maxim of thy will can always at the same time hold good as a principle of universal legislation.” It could be said that the formula of this imperative crystallizes the idea of a person as a rational being, whose behavior is guided not by self-interests, but by the requirements of what is universally right. The task of a rational being is to contribute to universal rationality.

This basic moral principle is available to every human person; it manifests itself as a “fact of reason” given directly, as a consciousness that makes itself felt as soon as the subject begins to think (5:31): “Pure reason is practical of itself alone and gives (to man) a universal law which we call the moral law.” This conclusion can be interpreted in the following way: the lawgiving reason is the carrier of rationality and at the same time is the source of motivation for disseminating this rationality. The realization of rationality happens through law: the rational being is supposed to enact a universal law; thanks to this law, reality would be transformed into a rationally ordered unity. The rational law grants unity to the surrounding world as well as to the human person.

Personal development consists in the more and more complete subordination of everything to the law. This idea, at least in Kant’s view, is synonymous with the idea of the realization of the highest good, constituting the conclusion and the message of his philosophy. However,

this idea also entails great difficulties. If we refer this message to the common moral intuition concerning human persons, the following thesis arises: the obvious conclusion of Kant's practical philosophy is the postulate of the moral self-improvement of the human person as a rational subject. The realization of the absolute good (i. e., the generation of a good will) can be achieved through the inner, individual effort of reason to subordinate the will completely to moral law, discovered in the sphere of thought. Essentially, good will consists in the ability to fulfill a moral duty without any regard for one's own inclinations, or even at the cost of the natural yearning for happiness. If we applied to this ability the traditional term "virtue", then virtue would by no means provide the human person with the possibility of acting effectively in the real world. This is where we come across a serious problem: the actual *helplessness of virtue*.

In his theoretical philosophy Kant has shown that reason is dialectical in nature (i. e., it inevitably faces problems it cannot resolve). The question of realizing the good turns out to be one such problem. But while the antinomies of theoretical reason (such as the question of the world's beginning in time or the question of the divisibility of matter) pose only intellectual problems and can essentially be justified by the limitations of human cognition, the antinomy of practical reason also has its "existential weight": it can shake our conviction about the absolute validity of the moral law and hence introduce a dangerous kind of chaos into the life of the individual and of the entire human community. This is why the discussion of the possibility of realizing the good takes on a particular significance.

In the *Critique of Practical Reason* Kant tries to prove the lack of contradiction, or even a kind of necessity, for the postulate that demands the reality of goodness and makes it hypothetically possible. As we know, according to his solution, the existence of specific causality between virtue and goodness can be consistently postulated by practical reason; this postulate is equivalent to assuming

the existence of a cause of all nature, distinct from nature itself and containing the principle of this connection. ... It follows that the postulate of the possibility of the highest derived good (the best world) is likewise the postulate of the reality of a highest original good, that is to say, of the existence of God.

It turns out that the elimination of the antinomy of practical reason, and thus the realistic expectation that the highest good will be realized, is

only possible on the assumption that God exists. Theoretical reason is not capable of determining whether God indeed exists, but practical reason can formulate the postulate of God's existence without contradiction. This opens up a new perspective, explored by Kant in his *Religion within the Limits of Reason Alone*. Let us take a closer look at this religious perspective.<sup>5</sup>

The first words in Kant's *Religion* – “the world plunges in evil” (6:19)—introduce quite a different tone than we can find his works on the philosophy of morality, where the concept of evil appears only occasionally. What caused such a surprising change? In the journal article “About the radical evil in human nature” that was then included in Kant's book on religion as its “First Piece”, his complex considerations can be virtually reduced to searching for the answer to a simple question: How is it possible that a rationally thinking person, sensitive to the moral law (hence having all possible tools to do good), is the instigator of evil that pervades the surrounding world? In spite of all the effort to account for it in a rational way Kant does not give a satisfactory answer to this question and admits that the essence and origin of evil are incomprehensible. It means that the irrationality of evil disturbs the rational structure of our thinking. This conclusion in the first part of Kant's *Religion* had to affect his attitude to ethics.

Kant regarded ethics as the best expression of rationality, the most perfect tool, when used in the right way, to make it possible for persons to reach their full potential. This excellence involves begetting good will in oneself—the only good “without qualification”. Kant, attracted by the rational “moral law in myself” seems to claim that to accomplish this good will—will that is totally submitted to reason—one needs nothing but reason. However, both his reflections on evil and his observation of others changed this optimistic view. With a certain astonishment Kant remarks that evil is not done exclusively by bad and immoral people. To see how much evil can be generated by the complexity of human relations we do not have to suppose that people are steeped in evil; on the contrary, it is enough to assume the presence of people—even people of good will can destroy each other's moral inclinations (6:94).

---

5 I write more about this problem in “*The relationship between ethics and religion in Kant's philosophy*”, Valerio Rohden (ed. et al.), *Proceedings of X. International Kant Congress* (Sao Paulo: Walter de Gruyter, 2005), vol.3, 53–62.

This leads Kant to formulate the thesis that clearly contradicts his earlier statements on good will: it turns out that even a person of good will—good will that was regarded earlier as the “good without qualification”—can be a tool in the hands of evil (6:97). This very consciousness of the moral law and eagerness to submit to the ensuing duties do not guarantee the accomplishment of good. Rational ethics seems to be helpless against evil: persons of good will cannot be sure their rational actions will not yield adverse effects that render a service to evil; evil in an incomprehensible way destroys the order established by the moral law. We can interpret this as follows: evil disturbs the unity of human personhood.

Is there a way out of this situation? Note that, according to Kant, mankind remains in the hands of evil because there is not “a unifying principle” among people. How should we understand this? To remedy helplessness in the face of evil, people of good will must combine their efforts and create a community (6:97). The obligation to join such a community is a special kind of duty. Duty, as it is generally known, is the key concept of Kant’s ethics. The most exhaustive list of duties is shown in his *Metaphysics of Morals* and it results explicitly from these considerations that a person has duties only toward oneself and other people. To do moral duty lies within a person’s power, whereas an obligation to create an ethical community is the task of humanity; on this a single person does not have sufficient impact. Thus, it is a special kind of duty—not of person to person, but of mankind to itself (6:97). To be able to fulfill this unusual duty an unusual assumption is necessary: the ethical community can be thought only as a religious community bound by God’s commandments (6:98–100). Kant expresses this in his conclusion that “ethics leads inevitably to religion” (6:6).

Talking about the religious community created not by mankind but by God, Kant underlines the limitations of human person. In moral terms: evil undermines rationality and the unity of the human person; the religion of reason, whose only substance is morality, makes this unity possible again.

## 37. How To Be a Good Person Who Does Bad Things

*Robert Gressis*

### 1. Introduction: Disposition vs. Propensities

In this essay I want to evaluate the relationship between the disposition and first-order maxims. More specifically, I want to examine how someone can be a good person while still sometimes doing immoral things.

In the *Religion* Kant introduces a pair of related terms that can easily be equated: “evil disposition” and “propensity to evil”. They are not, however, the same. The evil disposition is a noumenally adopted maxim (Kant says it is not adopted “in time” [6:25]), whose content is to subordinate the incentives provided by the moral law to those agitating in favor of advancing one’s self-love. Its contents can be loosely expressed as: when forced to choose between advancing my happiness and fulfilling my obligations, advance my happiness. Since it helps to explain why an agent adopts each maxim, but is not itself explained by any more fundamental maxim, it can be seen as an agent’s “supreme maxim”.

To understand the propensity to evil, one must know what a propensity is in general. Here is how Kant first defines a propensity, in the first edition (1793) of the *Religion*: “By *propensity* (*propensio*) I understand the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*), insofar as this possibility is contingent for humanity in general” (6:29). When he says a propensity is the subjective ground of an inclination, what Kant means is this: if you have a propensity to X, then that propensity allows for the emergence of an inclination to X. (He does not say, though, that if you have an inclination to X, it is because you have a propensity to X, except in the case of the propensity to evil.) So, if you have a propensity to evil, it is possible for you to develop an inclination to evil.

*Possibly* developing an inclination for something and *actually* developing that inclination, though, are different. How do you go from pos-



sibly developing something to actually developing it? Happily, Kant talks about this in the footnote to the above-quoted definition, appended in the *Religion's* second (1794) edition (6:29n):

*Propensity* is actually only the *predisposition* to desire an enjoyment which, when the subject has experienced it, arouses *inclination* to it. Thus all savages have a propensity for intoxicants; for although many of them have no acquaintance at all with intoxication, and hence absolutely no desire for the things that produce it, let them try these things but once, and there is aroused in them an almost inextinguishable desire for them.

If you have a propensity to X, then you can develop an inclination to X by “experiencing” or “trying” X. So, take the case of a propensity to intoxicants: if you have a propensity to intoxicants, then you can develop an inclination to intoxicants by experiencing or trying—i. e., consuming—an intoxicant.

So, to move from possibly having an inclination for X to actually having it, you have to experience X. Note that this allows us to talk about two states of the same propensity to X. First, there is the propensity to X *before* it is enlivened by an experience of X. I shall call a propensity in such a state an “inert” propensity to X. (Obviously, if your propensity to X is inert, you don’t have any inclination to X.) Second, there is the propensity to X *after* it has been enlivened by an experience of X. I call a propensity in this state a “lively” propensity to X. (Obviously, if you have a lively propensity to X, then you have an inclination to X.) Another feature of propensities is that if you have a propensity to X, then you are such that, when you experience/try X, not only do you develop an inclination to X, but you develop a *powerful, abiding* inclination to X.

Summing up these three characteristics of propensities: propensities can be either inert or lively; to move a propensity from inert to lively, you have to experience/try whatever that propensity is a propensity to; and lively propensities ground powerful, abiding inclinations. In the next section I shall apply these three characteristics of a propensity to the propensity to evil, and in the final section I shall show how having a propensity to evil enables even someone of a good disposition to act immorally. (I do not have space to explain how someone of an evil disposition can act from respect).

## 2. The Propensity to Evil

Like any propensity, a propensity to evil can be either inert or lively. For a propensity to evil to become lively, the person whose propensity it is must experience or try evil. If the person experiences/tries evil, his or her propensity to evil will ground a powerful, abiding inclination to evil. Two questions naturally arise: (1) what is it to experience/try evil? and (2) what is an inclination to evil?

I can think of only two candidates to fill the role of experiencing/trying evil. You experience/try evil by suffering it—that is, by being the victim of it. (The child molester molests her children because she was molested as a child.) Alternatively, you experience/try evil by engaging in it—that is, committing it. (The serial killer serially kills because he tortured animals when he was young.)

Kant could have meant either of these alternatives. Whichever one he means, though, a problem arises. Let us assume that by experiencing evil, he means suffering it. So, if you suffer evil, your propensity to evil will switch on. The problem is, Kant thinks people are morally responsible for their propensity to evil: “this propensity must itself be considered morally evil” (6:32). But if the only reason your propensity to evil is lively is that someone else enlivened it for you, then your propensity to evil cannot be seen as morally evil—that is, as something you are to be blamed for.

Assume, then, that you enliven your propensity to evil by perpetrating evil. The problem here is that you cannot perpetrate evil unless you are first inclined to do so. But you cannot be inclined to do so unless your propensity to evil is lively. The only way to enliven your propensity to evil is by perpetrating evil. This is a classic villainous-chicken-and-evil-egg problem.

To solve both of the above problems, we must introduce the evil disposition as responsible for the propensity to evil. Let us first look at how this move solves the second problem, the one that arises when one understands “experiencing/trying” evil as committing it. It is by choosing an evil disposition that you enliven your propensity to evil. Since this choice is noumenal, your propensity to evil does not need to be lively before you make it. In other words, a noumenal perpetration of evil enlivens an inert, phenomenal propensity to evil, and this then gives rise to a phenomenal inclination to evil.

Maybe you think that solution is too metaphysical. If so, here is a less metaphysical version of the solution, this time to the first problem,

the one that arose when I understood experiencing/trying evil as suffering it.

Your phenomenal, inert propensity to evil becomes lively when you suffer evil at the hands of someone else. Once it is lively, it gives rise to an inclination to something related to what you suffered (so, child abuse could give rise to one kind of inclination, namely the inclination to abuse children, but more likely it is going to give rise to some of a possible range of inclinations, such as wanting to hide from others, wanting to dominate others, etc.). However, this inclination is not yet one that you are morally responsible for: you are morally responsible for it only when you indulge it, because when you indulge it, you show yourself to identify with and support it. In this moment you bring upon yourself an evil disposition, and become responsible for your propensity to evil. (Note that it is not the other way around—your propensity to evil is not what is responsible for your free decision to identify with the inclinations it grounds.)

One last thing about the propensity to evil: it “cannot be eradicated [*ausgerottet*—literally, “rooted out”] (for the supreme maxim for that would have to be the maxim of the good, whereas in this propensity the maxim has been assumed to be evil)” (6:32). For the propensity to evil to be such that it could be “rooted out”, you would have had to have opted for the good disposition over the evil one (you would never have nourished it, and it would have withered away); if you had chosen a good disposition, you never would have enlivened your propensity to evil. Once you enliven it, though, you are stuck with it, even if you replace your evil disposition with a good one. With this in mind, let us now go back to the role of dispositions.

### 3. Maxims

I said earlier that the evil disposition is a supreme maxim that explains the selection of one’s lower-level maxims. How does it “explain” this selection, though? One possibility is that it could explain such selection through logical entailment. On this view, whenever you have a maxim, you just do whatever it is your maxim says you do. For instance, if my maxim ran, “when going to a movie theater, buy some popcorn”, then whenever I went to a movie theater I would try to buy some popcorn. If I went but nonetheless did not try to buy popcorn, then it would follow that the popcorn-buying maxim was not my maxim. So, on this

logical entailment view of maxims, you cannot have a maxim and even occasionally fail to act according to it.

Applying the foregoing to the evil disposition would mean that, if you had an evil disposition whereby the incentives of morality were to be subordinated to the incentives of happiness, then whenever you had to choose between happiness and morality, you would *always* choose happiness over morality. Otherwise, the evil disposition would not in fact be your maxim. On that reading, an evil person (i. e., someone with an evil disposition) would *never* prioritize respect over sensibility and would act from respect only when he did not have to give up *anything* to do so. Similarly, on this reading a good person would *never* subordinate morality to personal happiness.

There are a variety of reasons for thinking this is not Kant's view. First, he writes of evil people's behavior: "The human being is *evil*,' cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it" (6:32). If an evil person *always* subordinated morality to self-love, then one would have expected Kant to say such a person incorporated *frequent* deviation from the moral law into the supreme maxim, not just *occasional* deviation.

Second, the way Kant talks about grace in the *Religion* suggests that he believes evil people can sometimes act from respect for morality, even when, to do so, they have to sacrifice their sensible interests. "Granted that some supernatural cooperation is also needed to [the evil human being's] becoming good or better, whether this cooperation only consist in the diminution of obstacles or be also a positive assistance, the human being must nonetheless make himself antecedently worthy of receiving it" (6:44). In other words, you cannot become a good person without God's grace, but even with God's grace, you have to do something to receive it: you have to make yourself worthy of it.

Third, Kant is clear that good people can sometimes deviate from the moral law. To become a good person, you have to undergo a revolution in your moral disposition (with, as we have just seen, help from God). Even after you become good, though, your work is not yet over, for you must now begin the long process of gradual moral reform, and this process will doubtless be marked by a few missteps (6:46–7):

The restoration of the original predisposition to good in us is ... only the recovery of the *purity* of the law, as the supreme ground of all our maxims

... a human being, who incorporates this purity into his maxims, though on this account still not holy as such (for between maxim and deed there still is a wide gap), is nonetheless upon the road of endless progress toward holiness.

Fourth, in describing the three “grades” of the propensity to evil, Kant explains how a good person can sometimes go wrong: through frailty. Frailty is “the general weakness of the human heart in complying with the adopted maxims” (6:29), and if I am frail, “I incorporate the good (the law) into the maxim of my power of choice; but this good, which is an irresistible incentive objectively or ideally ... is subjectively ... the weaker (in comparison with inclination) whenever the maxim is to be followed” (6:29).

Finally, consider this remark Kant makes about maxims in general in his *Lectures on Pedagogy*: “The child should learn to act according to maxims whose fairness it itself understands” (9:480). If your maxims unfailingly determine what you do, then it makes no sense to say you have to learn how to act on them.

What is needed is another interpretation of maxims that makes these deviations understandable. So here goes. I think there are two kinds of maxims in Kant: “dispositional maxims” that we always act on, and “aspirational maxims” that we have to learn to act on. (I shall not say any more about aspirational maxims.)

Dispositional maxims are motivating judgments that serve as the major premises in practical syllogisms. Here is an example:<sup>1</sup>

Healthy things are good. (Maxim)

Exercising is healthy. (Rule of skill)

Therefore, exercising is good. (Conclusion that can serve as a maxim in another practical syllogism)

To have a dispositional maxim is to see something as good, and if you see something as good, you are to some extent motivated to do it.

You have a dispositional maxim associated with every kind of thing you see as good, so each of us holds a variety of dispositional maxims at any given time. Moreover, you can see the same kind of thing as both good and bad if you have two opposing dispositional maxims relating to it. For instance, I think exertion is bad but health is good, so I see ex-

---

1 I take this account of maxims, as well as this example, from Richard McCarty, “Maxims in Kant’s Practical Philosophy”, *Journal of the History of Philosophy* 44.1 (January 2006), 65–83, 69.

ercise as both good and bad. I see health as good but refraining from fattening food as bad, so I see dieting as both good and bad. As it turns out, every one of us accepts both the moral law and the law of self-love as maxims, so each of us sees doing our duty as good and promoting our happiness as good. If we accept this portrayal of maxims as dispositional maxims, then much of what Kant has to say about the good and evil disposition and their relationship to first-order maxims becomes more understandable.

First, let us call the moral law as we hold it the “Moral Maxim” and the law of self-love the “Prudential Maxim”. The Moral Maxim runs, “Doing your duty is good” and the Prudential Maxim goes, “Making yourself happy is good.” If you have an evil disposition, you subordinate the Moral Maxim to the Prudential Maxim, such that you see making yourself happy as *better than* (or, arguably, *just as good as*) doing your duty. If you have a good disposition, you subordinate the Prudential Maxim to the Moral Maxim, so that you see doing your duty as better than making yourself happy (see e.g., 6:30, 36, and 46).

Now the question arises, if you see X as better than Y, how can you ever choose Y? That is, perhaps you think my analysis of maxims does nothing to make it more understandable how a good agent can do bad or how an evil agent can do good.

But do not forget about the propensity to evil. The propensity to evil always remains with you, even if you choose no longer to identify yourself with it. For one thing, if you have a propensity to evil that you were at one point in league with, then you will most likely have developed habits before you revolted against it, and so you will sometimes thoughtlessly recapitulate your old ways.

However, the propensity to evil must stay with us as something more than just a set of bad habits. It would not take an eternity to overcome one’s bad habits, but if the propensity to evil remained just as a set of bad habits, that is what we would have to say: bad habits are ineradicable! That does not seem right for creatures like us, who are endowed with freedom.

Instead, I see the propensity to evil as a kind of force field of immorality. To understand what I mean, hearken back to respect, as articulated in the *Critique of Practical Reason*. According to the second *Critique*, respect works like this: you consider doing something immoral, and the mere consideration of that immoral course of action incites in you a striking down of your desire to undertake it. To put it metaphorically, when you realize that an inclination is an enemy of morality, the moral

law within you shoots a laser beam at it, weakening, though usually not utterly destroying its moving force. The fact that the moral law within you has this power awes you, and makes you respect it. So, out of respect for the moral law—that is, in deference to what most truly expresses your will—you may steer away from this enemy (though of course you may end up negotiating, and joining up with, it). What the propensity to evil does, though, is put a force field around at least some of your inclinations. Normally, the laser beam of respect severely weakens a foe; but when the propensity to evil operates, the enemy sometimes has its shields up, deflecting, slightly or wholly, the laser. Upon seeing this, you gain a reverse respect for that inclination. You could think, “hey, normally when I realize that a desire is universalizable, I don’t want to do it as much. But in this case, I still want to do it. So maybe there’s nothing wrong with this desire after all!” From this realization, a lot of mischief can arise.

PART VIII

Cultivating Personhood in  
Religion and Theology





## 38. Kant's Idea of Autonomy as the Basis for Schelling's Theology of Freedom

*Hans Feger*

### 1. Introduction

Kant's basic insight into the essence of human freedom is that its proper application also includes its thorough misuse, and that only his transcendental legal process of Critical ethics protects against this natural misuse of human freedom. To be confronted with the abyss of freedom—with the burden of responsibility—is certainly not, as false pathos often suggests, an act of liberation from obligations nor a license “to do as you please”; rather, it is a confrontation with the enigmatic side of modern subjectivity. Kant's definition of transcendental freedom as the “self-creation” of the actor clearly reveals the inner instability of human freedom. No cause antedates it, no choice precedes it; even the motivation that would be empirically or psychologically evoked, such as the will to preserve oneself, would not do justice to its absolute essence. Kant's concept of freedom that is linked to the idea of autonomy is based on a timeless, *transcendental act*. In Kant's view, an action is only free as a specific kind of causality: as a spontaneous, timeless causality born out of freedom. This already reveals what critics of this concept of freedom have always objected to: that Kant does not provide any theoretical proof of a transcendental freedom, since only something that is itself unrecognizable can serve him as a justifiable ground.<sup>1</sup> In this respect Schelling developed the Kantian metaphysics of freedom further.

---

1 Precisely this lack of provability also constitutes the strengths and the incontestability of the Kantian concept of freedom. It (1) releases Kant from the metaphysical burden of proof, but without thereby discrediting freedom or absolute spontaneity as a metaphysical category, (2) remains intact as a *conditio sine qua non*, since without the foundation of an absolute spontaneity, a moral law could never be obligatory for the individual, and moral phenomena such as conscience, remorse, and responsibility for one's own actions would be inexplicable, and finally, (3) is the core of the peaceable evocation of the third anti-

In his *Philosophical Inquiries into the Essence of Human Freedom* from 1809 Schelling demands that “freedom once be made the one and all of philosophy.”<sup>2</sup> He connected this to a speculation about the beginning of Nature: at its foundation the world, in order to be visible as Becoming, first is to be thought of as non-existent. This recourse to the ontological precondition of a beginning of Nature is not an unfounded speculation but the consequence of a radicalized understanding of freedom. Through such an understanding Schelling intends to expand the Kantian concept of freedom as self-determination by including freedom’s *real and living* possibilities. According to him, Kant’s purely formal concept of freedom remains incomplete and leaves us clueless “as soon as we want to go into the more precise and definite.”<sup>3</sup> Here Schelling bemoans the lack of a positive conception that would permit us to determine not only the moral position but also the place of individual existence that is to make such a moral position its own.

If one regards the concept of freedom with respect to its real and living possibilities, then one needs to make out not only the Good of rational morality but also Evil as that which resists reason and with it the unfounded independence from all predetermined instances as a *positive* core, as a *possibility* of corruption. The real and living concept of human freedom is not simply the capacity for Good and formally to be conceived as a demand for the correct relation to oneself, but rather “a capacity for Good *and* Evil.”<sup>4</sup> In *The Concept of Anxiety* Kierkegaard uses practically the same formulation, by the way, when he writes: “Only for freedom or in freedom is there a difference between Good and Evil, and this difference is not in the abstract but only in the concrete.”<sup>5</sup>

With this, a transformation of the transcendental, philosophically-grounded causality of freedom takes place within German Idealism. Critical philosophy’s concept of autonomy as represented in Kant’s con-

---

mony (B560–86) that produces the proof that such an absolute spontaneity can mutually exist with natural causality consistently and that, in principle, it does not have to cancel itself out: only as a free being can the human being transcend oneself as a natural being.

2 Schelling, *Philosophical Inquiries into the Essence of Human Freedom* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1983 [1809]), VII, 351.

3 Schelling, VII, 351.

4 Schelling, VII, 352.

5 Kierkegaard, *Concept of Anxiety*, ed. Hirsch/Gerdes, fourth edition (Gütersloh: Gütersloher Verlag, 1995), 114 f.

cept of spontaneity and Fichte's model of self-positing then loses its validity. Schelling assumes the "factuality of freedom"<sup>6</sup> and investigates its foundation within an Absolute that develops itself, without wanting to throw himself "in the arms of . . . necessity".<sup>7</sup> In his *Philosophical Inquiries*, Schelling investigates the formal concept of freedom in terms of its innermost preconditions. He is not concerned with a clarification of terms as to how freedom is possible; rather, he seeks to interpret the effects of freedom "in the totality of a scientific world view".<sup>8</sup> For this reason Schelling begins the essay by writing about the fact that freedom is an intuitive "feeling" for every human being. The reality of freedom in its immediate evidentiality needs to be spelled out in order to convey the philosophy of Nature and Freedom. That means: in order to reconcile the purely formal concept of freedom in Kant's and Fichte's idealism with the real and living concept of freedom of a "higher realism", freedom needs to be understood as a "capacity for Good and Evil"<sup>9</sup>—(indeed, as the capacity to be able to speak out *against* absolute reason too. Without the negativity of Evil, the capacity to make moral judgments would have no application in the real world. Kant's capacity of will in the form of a moral imperative is insufficient "in order to practice self-determination." Rather, the person "must be determined, clearly not through external forces that go against his nature, nor through internal compulsion whether it derives from accidental or empirical necessity . . . , rather in itself as its being, i. e. , his own nature must be his determination."<sup>10</sup> For Schelling, subjective self-determination is only then an act of freedom if it is to be found where Kant would never have placed it, in Nature, that is, if it is compatible with a subjective conception of *Nature* as a process that generates *itself*.

In the remainder of this essay I shall present the consequences of this expanded conception of freedom for the questions as to (1) its existential dimension, (2) its reintegration into theology, and finally (3) the possibility of its systematic examination.

---

6 Schelling, VII, 336.

7 Schelling, VII, 338.

8 Schelling, VII, 336.

9 Schelling, VII, 338.

10 Schelling, VII, 384.

## 2. The Existential Dimension of Schelling's Concept of Freedom

The basic conception Schelling employs to connect to and go beyond Kant's concept of freedom arises from the notion that Evil is a problem of the *realization* of freedom. In this context Schelling refers to Kant's late writings, his philosophy of religion. In those texts Schelling observes the tendency, contrary to the intention of Kant's moral philosophy, to posit for itself a *natural* basis in order to do justice to the phenomenon of Evil, this being a topic for religion. Kant, too, had refused to understand Evil as a positive force or to see it as a privation of the Good, as the tradition would have it. In his late text, *Religion within the Boundaries of Mere Reason*, a text Schelling explicitly refers to in his *Philosophical Inquiries*, Kant had labeled Evil *radical*, "because it corrupts the foundation of all maxims" (6:42) and it is based on the "subjective ground of the use of one's freedom in general" that "precedes all sensible deeds" (6:6). But instead of interpreting Evil simply as a problem of the corruption of maxims or of the inversion of principles of conduct (see 6:34), Schelling goes beyond this and seeks, more fundamentally, to explain Evil as arising in a freedom that as a *human* attribute becomes a capacity for Good *and* for Evil. Such Evil, however, in its potentiality cannot be constituted through an inversion or a corruption of the subjective basis for the acceptance of maxims (Kant), but rather through the "divisibility of principles".<sup>11</sup> On the other hand, Evil cannot exist *without* the Good, because otherwise Evil would take the place of Good and through "false imagination" be "conceived as real" by people.<sup>12</sup> Only when Evil as a dark principle that must always exist *in essence* is affirmed through an act of freedom—as if it were the Good—only then can we speak about Evil as such. In itself Evil is only a possibility provided for in the nature of human beings and one that should not develop an independent existence. The possibility of Evil in humans is *necessary* (for the Good), but the reality of Evil exists only as a *contingency* that depends on the free decision of humans. In essence Evil is to be understood as potentiality and as such in terms of a morally necessary condition for the realization of the Good. This theoretical basis makes it possible for Schelling to speak of a moral existence that accepts Evil as non-being at *the foundation of existence* and no longer binds Evil, exclusively,

---

11 Schelling, VII, 364.

12 Schelling, VII, 390.

to a moral decision. In terms of moral human existence this implies the claim that the Good that does not contain in itself the temptation for Evil is not the real, living Good.<sup>13</sup> Seen in this way, Evil consists of making absolute what in its essence is relative. Evil does not arise out of a singular principle but rather from the false combination of principles of existence.

Kant had established the turn to the Good as the moral decision of a person. The possibility of moral action for him is a necessary condition for duty. The certainty of moral necessity serves for Kant as a *ratio cognoscendi* of human freedom of the will. However, his solution to the problem of action is a thorny limit case for his moral philosophy. In his *Religion* Kant for instance writes (6:49 f): "How it is possible for a naturally evil man to make himself a good man wholly surpasses our comprehension; for how can a bad tree bring forth good fruit?" The solution to this dilemma is known: even in the most evil person a kernel of Goodness remains in the form of the moral law. The Moral law "imposes" itself "irresistibly" as a facticity of reason even on the worst person (6:33). Thus Evil is not constituted through the absence of the moral law but rather through affirming sensuality's incitement as a condition for following moral law.

In a sense Schelling turns this problematic around. Even in the best persons there must be a kernel of Evil, or more precisely, a potentiality for Evil so that they may act in a good way. The possibility of holding back Evil is what establishes the Good as the kind of Good that stands the test of reality. In terms of its *mere possibility*, Evil must be regarded as its own entity and not simply as the mere absence of the Good (*privatio boni*). In its potentially perverse independence Evil preserves and supports the Good. Whereas for Kant human freedom is the condition for the possibility *and* the reality of Evil, for Schelling human freedom is only responsible for the *reality* of Evil. Even following Kant it would be possible to claim that temptation, that is, the possibility of the corruption of freedom, is a constitutive element of freedom. Schelling, however, modifies this thought into the recognition that

regardless of this general necessity ... Evil is always a person's own choice; Evil as such cannot provide the reason, and every creature falls through its own guilt. But precisely how *the decision for Evil or for Good* is attained in

---

13 See Schelling, VII, 467.

every human being is a question still shrouded by darkness and seems to require an investigation of its own.<sup>14</sup>

In the context of the existential-ontological position of the human being, *indifference* between Evil and Good is no longer an option either for Schelling: the human being as such cannot remain undecided, cannot be a floating middle term, a “crisis” between God and the Devil. The human being

is placed on that peak where he possesses the source of self-movement to the Good and to Evil in equal parts: the bond of principles in him is not a necessary one, but a free one. He is at the crossroads; whatever he may choose, it will be his deed, but he can not remain in undecidedness.<sup>15</sup>

Schelling’s answer to the problem of Evil is not the decision *before* the deed but rather the *decided* deed. Once an act has been done, it has been done forever. To decide to be good out of freedom is a decision that is *always* accompanied by the temptation to do evil. Schelling calls this existentially fragile condition “the anxiety of life itself” that “drives people from the center in which they were created.”<sup>16</sup> If a person chooses Evil in the finitude of life, he or she is subject to forces of actions and temptations that he or she uses primarily to his or her own advantage. These are forces that arise out of the person’s “own will”. The claim of the Good that the person chooses in her or his condition as “spirit” comes at the price of the struggle *not to be guided first and foremost by the self-preservation of his or her own life*.

The problem of *moral* judgment arises here, since it is a limit case that allows us to study both sides of morality that come from the possibility of choice: the questions of morality with respect to its consequences and with respect to its justification. It is a limit case because a capacity for moral judgment clearly is not under the dominion of reason but rather under the dominion of the bearer of reason: in an individual’s reason and in that individual’s tendencies of conduct. At the individual level the decision is made as to how moral behavior that obeys the basic principles of general legislation pertains to one’s individual moral conduct toward oneself; this includes as well the question of the good

---

14 Schelling, VII, 381 f, my emphasis.

15 Schelling, VII, 374.

16 Schelling, VII, 381.

life.<sup>17</sup> Turning Kant's moral philosophy into an existential one brings me to the second part of my talk, to the problem of a re-theologization.

### 3. The Reintegration of Schelling's Concept of Freedom into Theology

Schelling added to his anthropological interpretation of Evil a theological dimension. In terms of theology, his interpretation gives him a surprising premise that leads to the famous distinction between principles pertaining to the "ground of existence" and the "existency" of every being—including God's being. God is relieved of responsibility for Evil precisely because human freedom includes the option of nothingness in the mere *possibility* of Evil; for God there is not *even* the possibility of Evil because in him both principles, in humans being different and therefore also divisible, have to be one and the same. In this perspective, the oneness of God with the world does not oblate the freedom of human beings, as Kant concluded from the metaphysical objections that arise from the assumption of a deterministic nature. Creation is not absolutely determined but is rather God's self-revelation in a free being. Thus it becomes obvious to downgrade the "ground of existence"—or *Nature in God*—in human beings to the "creature's own will"; against this "God as existing" has to assert himself as a "universal will". Just as Nature in God, as the ground or "fundament" of his existence, is "a being that is inseparable but yet distinguishable from God",<sup>18</sup> so too human beings are representatives of the initial nature or of the "Ur-will";<sup>19</sup> however, the personality of a human being is different and only relative to what it is at its fundament. On the one hand, the human being arises from that "which in God's self is not *Himself*"<sup>20</sup>—that is, from Nature in God—and thus is a *creaturely* entity. On the other hand, the human being in the world is a separated "spirit"—in other words, a reflected divine entity that *acts as substitute* for the God that sees himself in "his own likeness". This also means it is solely incumbent upon that eccentric freedom of human beings whether

---

17 For Kant, by contrast, the conscience is only one "sich selbst richtende Urteilskraft" (6:289).

18 Schelling, VII, 358.

19 Schelling, VII, 364.

20 Schelling, VII, 359.



such freedom serves a universal cohesion of the entities human beings themselves also belong to or, on the contrary, imposes its own will at the expense of universality. Confronted with this choice *the human being* takes on responsibility for creation. This relative autonomy toward the creator constitutes the precarious human freedom that gives rise to the possibility of Evil. In theological terms, the demand here is that the human being's own will remain subordinate to universal will as the latter's "instrument". Thus the human being might do good and reveal God instead of betraying this unity through Evil by turning the universal will into an instrument and misusing it for his own will. Human life as a finitude that is nonetheless *spirit* constitutes the impediment human beings need to overcome in order to bring themselves to do good.

The existential uncertainty of this process results from the fact that Schelling disconnects the notion of freedom from its entanglement with the autonomy of the will and grounds it in nature. The consequence of this, however, is that identity and freedom are then severed from one another: "Identity becomes an object of choice for a subjectivity that behaves freely towards it. Thus, identity is precarious."<sup>21</sup> In Schelling's formulation: "the subject can never grasp itself as that, which it is". In the act of willing *oneself* the human being does not love God. Every individual has aspires to a center, but this egoism is a bond that only serves self-preservation and suspends freedom. For Schelling, and later, particularly for Kierkegaard, "self-being", if it is made into "omnipotence" and does not remain degraded "to fundament, to organ", is only the attempt to want to be a single self desperately—a will whose paradoxical relation to the world is engulfing and destructive. If the difference could not be traced back to something that establishes it as a relation (God), then, from an existential point of view, despair, discord and illness would be the determining forces of life. Like Schelling, Kierkegaard also formulates the inverse structure of the self as a proportion that connects to unity in a negative relation of dependence: "despair is the failed relation in the relation of a synthesis that relates itself to itself."<sup>22</sup> Like Kierkegaard, Schelling calls this inversion of the unity of relation *illness*. Thus, Schelling's *Philosophical Inquiries* points the way and prepares the

---

21 Siegbert Peetz, *Die Freiheit im Wissen. Eine Untersuchung zu Schellings Konzept der Rationalität* (Frankfurt am Main: Klostermann, 1995), 224.

22 Sören Kierkegaard, *Die Krankheit zum Tode*, in *Gesammelte Werke*, 24. u. 25. Abteilung, Emanuel Hirsch and Hayo Gerdes (eds.), fourth edition (Gütersloh: Gütersloher Verlagshaus, 1992), 11.

ground for Kierkegaard's (sin-theological) idea of the self, grounded on the fundament of despair.

#### 4. The Possibility of a Systematic Examination of Schelling's Concept of Freedom

Within these parameters, Schelling's systematic aspiration—I shall argue in concluding—is ambivalent. Certainly, in God the fundament precedes existence, but likewise existence is the precondition for the fundament. What to an existing subject looks circular has to be a unity for an absolute being (apart from whom nothing can be) that is nevertheless a living entity—indeed, life itself. Taking up Jacobi's theory of conflict, Schelling intends to show that dependence cancels out neither independence nor freedom and that the “law of the ground . . . is as originary as that of identity.”<sup>23</sup> In the context of this critique—that is, the critique of those “abstract systems in which all personality is impossible in general”<sup>24</sup>—Schelling creates his system of a *panentheism* that claims to express in all its parts (true to the exegesis of the panentheistic axiom: everything is God) the fact of freedom as God's self-representation, as *Becoming in Becoming*. The narrative logic of this endeavor takes up the remembering repetition of the genesis of human freedom in order, by retheologizing it, to lead it beyond itself out of the interior spaces of religion within the boundaries of mere reason. It is crucial for Schelling's systematic approach that it does not result from a “relative independence from God” but rather from an “absolute independence from God”.<sup>25</sup>

---

23 Schelling, VII, 346.

24 Schelling, VII, 412.

25 Michael Theunissen, “Schellings anthropologischer Ansatz”, *Archiv für Geschichte der Philosophie* 47 (1965), 183. In evaluating the status of Schelling's anthropology, Theunissen comes to the judgment that he makes excessive demands on himself, because he does not take a stand on the problem of a derived absoluteness of human beings, and therefore does not develop anthropology in the sense of a *Prima Philosophia* or an anthropological idealism. In order to understand that the facticity of freedom differs radically from the natural, one “muß das Hauptproblem der schöpfungstheologischen Anthropologie die Frage sein, wie die Absolutheit, die in der Freiheit liegt, zugleich Nicht-Absolutheit, d. h. gesetzt sein kann” (179). But here Schelling's arguments fall apart (see his interpretation at VII, 354): “Schelling gibt seine Intention bei dem Versuch ihrer Ausarbeitung preis, indem er der Absolutheit der menschlichen Freiheit ein im Fortgang des Gedankens immer stärker werdendes Übergewicht

The two apparently contrary assertions, “There is system to divine understanding, but God himself is not a system but rather a life”,<sup>26</sup>—if one pulls them together—provide the ground for justifying an anthropological understanding of the concepts. The human soul “partakes in knowledge of Creation.”<sup>27</sup> For this reason Schelling infers in his *Weltalter*-philosophy: “It is clear that he who could write the story of his own life starting at its fundament would have epitomized, in so doing, the history of the universe.”<sup>28</sup> Much more could be said about all this, but for my present purposes this much will suffice.

To conclude, I would like to address two historical consequences of this understanding of system: Kierkegaard will turn around this problematic of an open system and radicalize Schelling’s conception of an existential freedom into an anti-systemic objection that puts an end to speculative idealism: existence “is itself a system—for God; but it cannot be a system for any particular existing spirit.”<sup>29</sup> Despite the theoretical proximity, Schelling remained for Kierkegaard a striking example of the fact that speculation is indifferent to existence, that in speculation “there is no result at all and no final decision.”<sup>30</sup> However, precisely this misinterpretation makes clear how Kierkegaard is beholden to Schelling’s point of departure: both focus on the question as to how a system of existence, inevitably, can also be thought of as a system of freedom.

Heidegger’s 1936 *Schelling Lecture*, on the other hand, flagrantly levels Schelling’s distinction between ground and existence. To put it

---

über die Deriviertheit gibt, bis schließlich die Deriviertheit ganz verschwindet” (180 f). In his *Weltaltern* (1813) Jochem Henningfeld grants the status of a “Basis metaphysischer Spekulation” to one of Schelling’s anthropological reflections. “Der scheinbar hybride Versuch, Gott und das Universum vor unseren Augen entstehen zu lassen, hat ein anthropologisches Fundament. Die Anthropologie als Wissen vom Wesen des Menschen ist hier kein untergeordneter Teil des Systems, sondern tragende Stütze für den Aufbau des Systems” See Jochem Henningfeld, *Die Menschlichkeit des Absoluten*, in *Philosophische Anthropologie im 19. Jahrhundert*, ed. F. Decher and J. Henningfeld (Würzburg: 1991), 37–49.

26 Schelling, VII, 399.

27 Schelling, *Die Weltalter*, 4.

28 Schelling, VIII, 207. On this point see the thesis of Wolfgang Wieland, that Schelling’s myth of the becoming God is his “Auslegung der menschlichen Selbsterfahrung” (W. Wieland: *Schellings Lehre von der Zeit. Grundlagen und Voraussetzungen der Weltalterphilosophie* [Heidelberg: 1956], 77).

29 Sören Kierkegaard, *Abschließende unwissenschaftliche Nachschrift I*, 111.

30 Sören Kierkegaard, *Abschließende unwissenschaftliche Nachschrift I*, 246.

briefly, the introduction of Evil into the metaphysical system dissolves systematic thought as such and in its place emerges the metaphysics of decision, where everything that is called "being" boils down to *how* human beings exist. Heidegger reinterprets Schelling from the perspective of his own philosophy when he writes: only because "the possibility and the reality of Evil as a finality of free human existence metaphysically reach into the absolute ... is human freedom at all able to make a justified claim to the basic character of a central point of the system."<sup>31</sup> This quotation shows the terminological weakness of the reference to Schelling, because if cited correctly, according to Schelling's *Edited Works*,<sup>32</sup> only the *reality* of Evil could be counted as finality of free human existence. As a *possibility* Evil is a given in the nature of human beings and is necessary for the decision for the Good. In other words, it is in no way justifiable to claim, as Heidegger does in the *Schelling Lecture* and in later seminars on the same topic, that the introduction of Evil into the metaphysical system requires us to abandon systematic thought itself.<sup>33</sup> The "decisive estimation of the Good" is in no way, as Heidegger would have us believe, the fatal "claim of Evil".<sup>34</sup> In

31 Martin Heidegger, *Schellings Abhandlung über das Wesen der menschlichen Freiheit (1809)* (Tübingen: 1971), 191.

32 Heidegger refers in his Schelling-interpretation to the edition: F. W. J. Schelling, *Das Wesen der menschlichen Freiheit*, ed. C. Herrmann (Leipzig: Meiner, 1925). The quote from Schelling is wrongly shown there (instead of "die Entscheidung für das Böse und Gute" it should correctly read "die Entscheidung für Böses oder Gutes" (VII, 382). This error was transferred into modern Schelling editions, including even that by Horst Fuhrmans (Stuttgart: Reclam, 1964).

33 See the passage where the lecture becomes Heidegger's self-interpretation: "Wer den Grund dieses Scheiterns [Schellings] wahrhaft wüßte und wissend bewältigte, müßte zum Gründer des neuen Anfangs der abendländischen Philosophie werden" (4).

34 The central passage in Heidegger's *Schelling-Vorlesung* reads: "Die menschliche Freiheit—hieß es—ist das Vermögen des Guten *und* des Bösen. Vielleicht haben wir bisher gar nicht recht beachtet, daß Schelling sagt: zum Guten und zum Bösen; oder wir haben es höchstens insoweit beachtet, daß wir im Stillen an dieser Fassung einen Anstoß nahmen als einer unscharfen. Denn es müßte doch, streng genommen, heißen: zum Guten *oder* Bösen. Nein; solange wir dies meinen, haben wir die vorgegebene Wesensauslegung der menschlichen Freiheit noch nicht gefaßt. Denn die Freiheit als wirkliches Vermögen, d. i. entschiedenes Mögen des Guten ist in sich zugleich auch das Setzen des Bösen. Denn, was wäre ein Gutes, das nicht das Böse gesetzt und übernommen hätte, um es in die Überwindung und Bändigung zu bringen? Was wäre ein Böses, das nicht in sich die ganze Schärfe eines Widersachers des Guten entwickelte? Menschliches Freisein ist nicht die Entschiedenheit zum Guten

order to do justice to Schelling's system of freedom, we need to return to Kant's reflections on *radical Evil* in his *Religion within the Boundaries of Mere Reason* and to interpret Schelling's notion of freedom in the *Philosophical Inquiries* as a practical surpassing of it.

---

oder zum Bösen, sondern die Entschiedenheit zum Guten und zum Bösen oder die Entschiedenheit zum Bösen und zum Guten. Nur dieses Freisein bringt je den Menschen in den Grund seines Daseins, so zwar, daß es ihn zugleich heraustreten läßt in die Einheit des in ihm ergriffenen Willens zum Wesen und Unwesen. Dieser ergriffene Wille ist Geist und als solcher Geist Geschichte" (Heidegger, *Schellings Abhandlung*, 188).

## 39. Moral Theology or Theological Morality?

*Mohammad Raayat Jahromi*

### 1. Introduction

Kant, undoubtedly, was a great philosopher, who created a system of thought that most subsequent philosophical movements can be traced back to. Thinking deeply on the thought structures of continental and analytical philosophies, one finally concludes that Kant's thoughts are significantly present in contemporary movements. His pre-Critical and Critical works follow one single and common guideline, signifying coherence among the elements of Kant's thought. Thus, it can be assumed that his "Optimism" essay (1759) is an introduction to his *Critique of Pure Reason* (1781), and *Religion within the Boundaries of Mere Reason* (1793) is a consequence of his second *Critique* (1788).

Although the epistemological approach and the birth of Critical philosophy were the fruits of Kant's intelligence, his critical attitude toward antecedent thinkers and his specific positions relative to the scientific atmosphere and philosophical schools of the time had a notable influence on forming the Kantian school of thought. Moreover, politics, society, religion, and culture played major roles in his thinking and unveil the way Kantian philosophy combined with the Enlightenment age: Kant was a thinker of the eighteenth century and surely held the ideas of his time. The first core of Critical philosophy was formed in the pre-Critical thesis "Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality" (1764) and was rooted in Kant's sensitivity toward the scientific and philosophical procedure of his time, controlled by Newton and Hume.

The pre-1781 Kant, or better to say, the hibernating in dogmatism Kant, was under the influence of Newtonian physics. He thus employed pure reason to justify Newtonian physics, and in the named work he insisted on establishing natural theology based on philosophical necessity. Being drowned in the waves of the Newtonian arena, until his 1770 Inaugural Dissertation, "On the Form and Principles of the Sensible and Intelligible World". Kant philosophized without any consideration of

Newtonian Physics' weaknesses and limitations. However, as he discovered the Critical philosophy, following Hume's shock, he changed his previous attitudes and concluded that pure reason can be as fragile as Newtonian physics. In this way, Kant's three *Critiques* were born under the light of a Copernican revolution, the epistemological position of criticizing the powers of human reason.

Kant's pre-Critical mind was haunted by many doubts and questions. What features have made Newtonian physics publically favorable and unquestionably acceptable? Can we trace its acceptability back to social and cultural reasons? What happened to metaphysics? What causes the decline in morality? Is it justifiable to say that developments in mathematics and natural sciences make metaphysics unnecessary? The pre-Critical Kant, or the dogmatic Kant, responded to these questions by considering the principles of Newtonian physics. Its principled nature reached a point that risked labeling metaphysics and morality as "unscientific" and "logically unexplainable", isolating and demolishing them. Eighteenth century thinkers went so far that if they could not explain a matter by reason, they considered it illogical; they did not consider this as a proof of the weakness of reason. Meanwhile, the vague and futile arguments of the supporters of metaphysics intensified such positions. However, Kant never accepted attacking morality for not being scientific.

Along with arguments over Newtonian physics, eighteenth century European thought held another characteristic: the opposition between Rationalism and Empiricism. The former elaborated how to know objective realities, the latter illustrated how to identify the things. These two major realms of philosophy were categorized and developed under the important name of "epistemology". Kant's philosophizing in the pre-Critical era and Hume's empiricism, with the scent of a common doubt, made the situation much more complicated. The world of philosophy was too far removed from Kant's awakening and its manifestation through the Copernicus revolution.

At the time, the Rationalists insisted on innate suppositions, while the Empiricists relied on *a posteriori* and sensible suppositions. Hume was the one who preceded Kant in developing Skepticism, just as Montaigne did before Descartes. Nevertheless, the appearance of Critical philosophy was due not just to the dominance of Newtonian physics, or the poor condition of metaphysics, or the opposition of Rationalism and Empiricism. Kant is the philosopher of Enlightenment and it is unfair to limit his broad worldview to specific theories and to concentrate

just on his first *Critique*. Therefore, rethinking the cultural and historical background of Kant's system of thought is of great importance to us.

Another significant factor in forming the very essence of Kantian thought was the late seventeenth-century movement in the Lutheran church of Prussia. Kant's Pietistic background is clearly notable in his later works. In order to attain faith, Pietism considered reason as not enough but insisted on a pure heart and mind, and piety. It may be his childhood background that made Kant claim reason as not sufficient for knowing truths as God does, and as he says, made him put aside knowledge to make room for faith. Protestantism had a great influence on Kant's worldview, though he reinterpreted Protestantism and criticized the philosophical theology of Catholicism as well. Luther also affected Kant. Luther believed mankind's moral consciousness is the Supreme Court and highest criterion to identify right and wrong. He also noted that morality is a personal and individual matter that cannot be gained through reason, but only through faith.

Empiricism and Rationalism brought Kant the criteria of scientific objectivity in the form of synthetic *a priori* propositions. In Kant's philosophy, the main aim is to identify the *a priori* elements and apply them to experimental data. His humanity-centered metaphysics elaborated these propositions, confirmed mathematics and physics, and challenged the objectivity of traditional metaphysics. By increasing the authenticity of mathematics and physics and better understanding Hume's doubts, Kant extended Hume's doubts from causality to the entire realm of metaphysics, rejecting metaphysical claims to attaining objective knowledge.

Here Kant felt the risk of demolishing religious elements and tried using morality to set religion free from theoretical reason and to protect science from doubts. He adopts the self-criticism method to assert that accepting its absolute freedom is the only appropriate subject for reason. Kant's *Critiques* were the fruit of such an approach: his two-dimensions of humanity arose from his criticism of experimental and theoretical reason, and from his effort to prove the possibility of Newtonian physics and of moral duty. Kant believed religion should be interpreted in the realm of morality; this idea was illustrated in "On the miscarriage of all philosophical trials in Theodicy" and *Religion within the Boundaries of Mere Reason*.

Thus, Kant eliminated theoretical reason from metaphysics and took a new position, in order to save metaphysics from destruction. The former concerns of metaphysics are now principal belongings of morality,



and should be handled by practical reason. Morality serves as the root of faith and gives meaning to the concept of God. Kant tried to found faith on morality, for in his system of thought theoretical reason has nothing to do with the original and principal religion. The original religion, in Kant's idea, is the moral religion and can be imagined solely through the dimensions of human reason. He calls the moral religion the "religion of good life conduct" (6:51).

## 2. Religion and the Ethical Commonwealth

Kant looks for *Religion within the Boundaries of Mere Reason*, as the title of his most important thesis on religion reveals, and tries to establish a morality-based religion. True religious concerns have nothing to do with theoretical reason, and Kant prepares the basis for denying knowledge in his *Critique of Pure Reason*, so that room can be made for faith in the realm of practical reason. His *Religion* presents this approach.

Kant based religion on two different attitudes, making it possible to distinguish two categories of concepts from each other. These concepts present basic ideas and terms in Kant's language, and thus, help him illustrate his ideas on religion and its connection with morality. He groups historical religion, revealed religion, the religion of rites and rituals, faiths of divine worship, and ecclesiastical faith into the first category, while putting moral religion, rational religion, morality reforming religion, true religion, pure moral faith, the true church, the visible church, the ethical state of God, the ethical commonwealth, and the universal religion of reason in the other. Hence, he criticizes the concepts mentioned in the first category, and strengthens the basis of the ones in the second category.

Kant defines religion as "the recognition of all our duties as divine commands" (6:154). He believes such a definition will rule out the possibility of assertorical knowledge of God, and since our perception of super-sensual phenomena is not authentic, theoretical reason is not permitted to enter the realm of religion. Thus, all kinds of theoretical considerations in religion, manifested in the form of recognition of God, would be false and would end up in hypocrisy. Kant elaborates this idea in his thesis "On the miscarriage of all philosophical trials in Theodicy", considering the biblical story of Job and his friends.

Kant believed all religions are either "endeavors to win favor" or "religion of good life conduct" (6:51): the former is the religion of wor-

ship and prayer; the latter, the religion of morality. Historical religion, manifested in the faith of the Church, belongs to the first category (i. e., being subject to God's grace is no matter of specific deeds by human beings). On the contrary, the religion of morality finds its meaning in the context of actions and deeds (i. e., taking action in order to be good and live honorably). Hence, in Kant's view human beings can think of God's grace only when they have done all they can to become better. He says (6:96): "It is not essential, and hence not necessary, that every human being know what God does, or has done, for his salvation, but it is essential to know *what a human being has to do himself* in order to become worthy of his assistance."

Kant portrays human life as the permanent battlefield of good and evil inevitably fighting to knock each other down. He says the good can rule over a person's life if the person tries to fulfill his religious practices, finally resulting in moral perfection. According to Kant, the good is rooted in the morally legislative reason and a model of the good cannot be seen in experiential examples, for the model is *a priori* and has been placed in our reason beforehand. "And the required prototype always resides only in reason, since outer experience yields no example adequate to the idea" (6:105).

"This is a change of heart which must itself be possible because it is a duty" (6:108). On the other hand, the principle of goodness does not limit itself to any time (i. e., the primary elements of goodness have been placed in human nature since the very beginning of creation). Kant indicates that the principle of goodness has been manifested in Jesus Christ, who chose "to die to everything that holds [people] fettered to earthly life to the detriment of morality" (6:82). To him, Jesus is an example to the whole mankind and the true teacher of morality; in the same way, through accepting him, his followers are God's children. Jesus looks for those among them who acknowledge good deeds.

Accordingly, the true religion for Kant is the moral religion, or the religion of good life conduct. However, he says (6:121–2): "We should not ascribe to this good any other distinguishing trait except that of a well-ordered conduct of life ... There is absolutely no salvation for human beings except in the innermost adoption of genuine moral principles in their disposition." Thus, moral duties, viewed as divine orders, are the essence of religion. Nevertheless, Kant believes the ideas mentioned in historical religion, as opposed to moral religion, cannot be perceived by theoretical reason. "Nor, in general, can anything super-

natural be known to us, because all use of reason ceases precisely with it. For it is impossible to make these effects *theoretically* cognizable ... because our use of the concept of cause and effect cannot be extended beyond the objects of experience” (6:96). Therefore, Kant concludes that religion is a part of morality and faith finds its meaning just in the context of practical reason, not through the theoretical reason that leads to impossibly perceivable mysteries. On the one hand, according to practical reason, there is no mystery; we must take action and live a better life, and fulfill the *a priori* principles of practical reason (i. e., do our duties). On the other hand, historical religion believes in a divine mystery, but we cannot perceive it, so attempting to perceive it is totally futile.

Kant considers historical religion to be limited to the conditions of time and place, and thus subject to frequent changes. However, moral religion is firm and stable, rooted in the *a priori* principles of practical reason. Thus, the historical or experiential conditions do not intrude on it whatsoever (6:136–7): “For it is a plain rational faith which can be convincingly communicated to everyone, whereas a historical faith, merely based on facts, can extend its influence no further than the tidings relevant to a judgment on its credibility can reach.”

Criticizing the faith of the church, Kant notes that humanity is not able to affect the deity, for human power is limited to earthly creatures. Therefore, contrary to what has been mentioned in the faith of church, it is impossible to serve God unless human beings fulfill their duties toward themselves and others (6:137):

It does not enter their heads that, whenever they fulfill their duties toward human beings (themselves and others), by that very fact they also conform to God’s commands, hence, that in all their doings and non-doings, so far as these have reference to morality, they are *constantly in the service of God*.

Kant thus concludes that we should pass over church faith and describes the real kingdom of God, or the real church, in light of rational or moral religion. In fact, the ethical commonwealth would be a church where people try to live the right way and receive divine grace through living the right way and fulfilling their duties. According to his own epistemological basis, mentioned in *Critique of Pure Reason* (where he classifies twelve categories under four headings), Kant sums up the necessities and characteristics of the true church in the following four elements: (1) Universality, whence its numerical unity; (2) Quality (i. e., purity); (3) Relation under the principle of freedom; and 4) Modality, the unchangeableness of its constitution (6:101–2).

Here, the Kantian definition of religion comes to mind: to recognize all duties and responsibilities as divine rules. Accordingly, church faith can be distinguished from pure religious faith, and the definition of “serving God” can be reinterpreted (6:138): “For in pure religious faith it all comes down to what constitutes the matter of the veneration of God, namely the observance in moral disposition of all duties as his commands.” Thus, being members of an ethical commonwealth, we are citizens of a divine state on earth. Kant differentiates between religion and faith and believes there is just one true religion, the moral religion; Christians, Jews, and followers of other religions have tended toward a faith, not a religion.

According to Kant, historical faith and ecclesiastical faith are properly regarded as means of fulfilling pure religious faith, the real goal of religion; if we consider historical faith as the goal, not the means, the fulfillment of moral religion or pure rational religion will be postponed. In other words, the only fruit of holding exclusively to the ecclesiastical faith is the postponement of good behavior or moral lifestyle. Therefore, we need to make every effort for a gradual move from the ecclesiastical faith to the absolute authority of pure religious belief. This can be achieved only through the fulfillment of our duties and a moral lifestyle. We can then hope to establish the divine state or the moral state of God on earth. Hence (6:135): “The true (visible) church is one that displays the (moral) Kingdom of God on earth inasmuch as the latter can be realized through human beings.”

Kant states that the fulfillment of such an ideal is the first step to establish permanent peace in the world (6:153):

Such is therefore the work of the good principle—unnoticed to human eye yet constantly advancing—in erecting a power and a kingdom for itself within the human race, in the form of a community according to the laws of virtue that proclaims the victory over evil and, under its domain, assures the world of an eternal peace.

As we have seen, Kant holds a paradoxical attitude toward historical faith. He sometimes calls it inefficient and futile, yet elsewhere calls it a means of fulfilling the pure religious faith. He also distinguishes the natural juridical state from the natural moral state. He writes (6:130): “A *Juridico-civil* (political) *state* is the relation of human beings to each other inasmuch as they stand jointly under *public juridical laws* (which are all coercive laws). An ethico-civil state is one in which they are united under laws without being coerced, i. e., under *laws of virtue* alone.

In this way, we have four states: (1) the natural juridical state, (2) the natural moral state, (3) the juridical civil state, and (4) the ethical civil state. Kant concludes that human beings have always been pursuing a historical procedure from the natural juridical state to the ethical civil state (i. e., we have always tried to fulfill the principle of goodness through the establishment of an ethical commonwealth that needed to consider an ethical legislator named God). In such a society, the visible church is in charge of executing the rational laws authorized by the ethical legislator, who is ethically sacred, benevolent, and just. Hence (6:166): “The threefold quality of the moral head of the human race, which in a juridico-civil state must of necessity be distributed among three different subjects, can be thought as united in one and the same being.” The legislative force incarnates Divine Sacredness, the executive force shows Divine Justice, and the executive force visualizes Divine Benevolence.

### 3. Moral Theology

Considering elaboration and support of pure religious considerations as the major approach of theology, we can define two types of theology, based on theoretical and practical reason. One theology utilizes theoretical reason to confirm the major principles of religion, while the other tries to find those principles in the realm of practical reason. In other words, rational theology strengthens religion through theoretical reason. Yet, the theology based on practical reason criticizes rationalism in matters of faith and suggests perceiving religion through morality. Aquinas is the most distinguished supporter of rational theology. Kant, on the contrary, tries to make room for faith within the realm of morality and speaks of a theology of practical reason through a denial of theoretical reason.

“*Summa Theologica*” is the famous work of Saint Thomas Aquinas in natural theology. In Western philosophy, natural theology has combined Christian faith with philosophical attitudes of Greek, Roman, Jewish, and Muslim philosophers. Thus, Aquinas’ approach to the relation between religion and philosophy, or faith and reason, was influenced by all of these together. He uses Greek rationalism to confirm the Catholic fundamentals, and presents philosophy as serving religion; as such, he can be considered as one of the founders of Christian theology.

Those who call philosophy dependent on religion have tried to use it to form religious concepts, and to explain them to others who are not their coreligionists, or to defend their faith. During the Middle Age and the sovereignty of religion, the church made every effort to defend and develop Christianity and utilized philosophy to serve this aim. Philosophy served religion and theological schools were founded.<sup>1</sup>

The five demonstrations Aquinas mentioned in his book were an effort by theoretical reason to prove God's existence and confirm the Catholic Fundamentals.

After Thomas Aquinas, the fourteenth-century Christian philosopher, William of Ockham separated religion and philosophy, and criticized the fundamentals of Aquinas' theology. He believed theology cannot be considered as a science and called it a meaningless subject that natural reason has no access to. His impression on theology, cutting off the relation between metaphysics and the sacred teachings, remained for two centuries after him. "He stated that there is no way to perceive the concept of God through reasoning and demonstrations."<sup>2</sup>

Although Kant holds a position fundamentally different from Aquinas' philosophical theology, he shares some common points with William of Ockham. Kant criticized the rational theology of Aquinas as well, but in his own special interpretations and attitudes. Kant did not accept demonstrations based on theoretical reason to prove God's existence, and as he said, he put aside knowledge to make room for faith. However, Kant's idea of rejecting theology based on theoretical reason and any kind of rationalism in religious affairs is not ever a sign of Fideism, for his attitude is just an epistemological way of criticizing mankind's power for saving metaphysics.

We previously mentioned that God and matters concerning the divine had been the main philosophical topic considered by theoretical reason. Kant categorizes the proofs of God's existence into three groups: natural and teleological, cosmological, and ontological. He criticizes them and finally accepts moral theology based on practical reason in order to refrain from following Aquinas and his theology based on reason.

---

1 Mohammad Ilkhani, *Seven Sky (Haft Aseman)*, in Persian (Tehran: Religion and Philosophy, 2002), 89–91, my translation.

2 Mohammad Ilkhani, *The History of Middle Age Philosophy* (Tarikhe Falsafe dar Ghorroone Vosta) (Tehran: Samt Press, 2003), 537, my translation.

Another point is that Kant rejects the possibility of the three demonstrations based on his first *Critique* principles, especially the significant distinction between noumenon and phenomenon, limiting epistemology to the world of phenomena. Thus, one cannot prove God's existence with the aid of the empirical and causal world.

Kantian rational theology is rooted in practical reason and his basis of faith is morality. He says God's existence finds its meaning through morality, for it is an essential presumption for morality. Accordingly, Kant's critique of Aquinas' rational theology and rational demonstrations for God's existence does not indicate his denial of a divine entity. Kantian rational theology, based on practical reason, replaces Aquinas' theology based on theoretical reason. In summary, Aquinas tried to prove God's existence through theoretical reason and the empirical and causal world, while Kant portrays God as an idea of practical reason. Aquinas reaches God from the world of existence; Kant finds God in morality.

Kant believes that God is an essential assumption for morality, while Aquinas calls God an essential entity. In other words, according to Catholic fundamentals and the spirit of Middle Ages, Aquinas utilizes theoretical reason to present demonstrations for God's existence and makes God the most central subject in philosophy. Nevertheless, Kant believed that human beings do not need a transcendent Being to fulfill their moral duties. He appreciates free will so highly that he even tries to find the root of his moral system and God's existence in free will: the morality-supporter Kant, who tries to find religion in morality, considers God as an essential presumption. In Kant's idea, "the greatest mistake of former philosophers was to establish morality based on the divine order, for this amounts to heteronomy."<sup>3</sup>

Aquinas widely acknowledged revelation and always considered it as a bridge to the unknown world.

In general, he divides theology into two types. Revealed theology, which is directly derived from the Holy Bible and based on Christian faith, and natural or philosophical theology, which is created by man's natural reason. In philosophical theology, he considers reason as the organizer of religious concepts.<sup>4</sup>

---

3 Stephan Körner, *Kant*, (Harmondsworth, U. K.: Penguin, 1984), ch. 4, section 7.

4 Ilkhani, *The History of Middle Age Philosophy*, 403.

Meanwhile, Kant calls theoretical reason inefficient in the elaboration of religious concepts.

Furthermore, Kant considers only the moral teachings of the Holy Bible (i. e., teachings that support moral religion). However, Aquinas had believed the entire teachings of the Holy Bible to be worth reading and practicing, for the Bible comprises the principles of Christianity. He had been trying to prove and elaborate Catholic teachings through reason. We should remember that Kant was a moral Protestant who was raised in a pietistic family and educated during the Age of Enlightenment, while on the other hand, Aquinas was a Medieval thinker whose approach to rational theology was meant to strengthen Catholic faith.

As mentioned above, Kant was a Pietist, as is evident through his personal lifestyle and his works, especially his *Religion*. By contrast, ecclesiastical rites and rituals were of great importance to Aquinas, for they are basic principles of Catholicism. Another point is that, Kant explicitly rejects historical religion and sought to base religion on morality.

Aquinas refers to the essence of the existing world to prove God's existence; by contrast, Kant considers God in the realm of morality. Therefore, Kant and Aquinas are both foundationalists with regard to God's existence, searching for a foundation for belief in God. As a pre-modern foundationalist, Aquinas based Christian theology on reason. Kant, as a modern foundationalist, criticizes such demonstrations and refers to practical reason to strengthen religion. He believes God rules over the world through moral laws, and thus provides eternal happiness to those who fulfill their duties. According to Kant, morality tells us about God, an entity who is morally perfect, and denial of God is rejection of the moral nature of human beings. On the contrary, Aquinas says goodness is a divine attribute and God is the origin of all moral perfections. Goodness exists because God is the first and original reason for goodness. In short, Kant talks about a God who is proved through morality, while Aquinas calls the goodness of morality the result of God's existence, believing that since God exists, goodness also exists as a result, so that morality is dependent on God, not God on morality.

Contrasting his view with that of Aquinas shows that Kant distinguishes moral theology from theological morality. He rejects theological morality, for morality should not be based on theology but needs principles that establish the basis of goodness for us. In other words, the tendency to believe in God's existence comes from our belief and attachment to morality. While in the pre-modern era, prior to Deism, all af-



fairs including morality were based on the concept of God, such a feature drastically changed in Kantian attitude. As mentioned, there are four types of relation between historical and moral faith, theoretical and practical reason. The fourth type insisted on the denial of any rationalism in the realm of faith. This does not mean that Kant is a Fideist: he separated faith from theoretical reason and constituted it in the realm of practical reason (i. e., morality). He went further and based faith on practical reason. In fact, Kant opposes theoretical rationalism in the realm of historical faith on the one hand, and defends practical rationalism in the realm of moral faith on the other.

#### 4. Conclusion

Kant distinguishes between the natural juridical state and the natural moral state. Inspired by Hobbes' *Leviathan*, Kant interprets the natural juridical state as the battle between two persons and the natural moral state as the permanent battle between good and evil inside an individual, as previously mentioned. Kant considers the lack of a unifying principle in the natural moral state would lead to disharmony, for he believes that even people who have a good will destroy each other's moral nature. In his idea, the presence of a transcendent Being named God is necessary to unite human beings. He stresses that in order to achieve an ethical commonwealth we should leave the natural moral state and establish a united society of people (6:133):

... this highest moral good will not be brought about solely through the striving of one individual person for his own moral perfection but requires rather a union of such persons into a whole toward that very end, [i.e.] toward a system of well-disposed human beings in which, and through the unity of which alone, the highest moral good can come to pass.

He believes the principle of goodness can conquer the principle of evil only if a society is established where everyone fulfills their duty according to rational rules. In such a society humanity is the center of the world and duties have found new meanings: individual duties have been minimized and replaced with the social duty of each person toward his or her fellow human being (6:130): "An association of human beings merely under the laws of virtue, ruled by this idea, can be called an *ethical* and so far as these laws are public, an *ethico-civil* (in contrast to a *Juridico-civil*) society, or an *ethical community*." Kant notes that this society is

the utopia for everyone, as the moral duties in such a society are divine orders and its citizens are people of God who seek virtue under the moral government of God.

As a conclusion, we can restate the whole essay in terms of religious and theological attitudes. In the realm of religion Kant differentiates moral and historical religion. He says historical religion includes mysteries that cannot be solved or perceived by theoretical reason. In defining moral theology he distinguishes theoretical and practical reason. He calls theoretical reason incapable of perceiving religious truths, leaving them to practical reason. He defines moral religion as the recognition of all duties as divine rules; thus, the true religion is the moral religion. The essence of religion should be found in the *a priori* principles of practical reason rather than in mysteries that form the basis of historical religion.

In general, if we consider reason as the Greek *ratio* and interpret faith as attachment to the deity, then we can observe four types of relation between religion and philosophy, or faith and reason, in the Christian West: (1) philosophy as religion, (2) philosophy as a criterion to perceive religion, (3) philosophy as a concept to serve religion, and (4) philosophy as a concept distinguishable from religion. The philosophical theology of Saint Thomas Aquinas is the result of his perception of philosophy as a concept to serve religion. But in my opinion, Kant cannot be placed in any of the above types, because the basis for this division is the comparison between faith and theoretical reason; he separated faith from theoretical reason and constituted it in the realm of practical reason (i. e., morality). As we stated before, Kant is not a Fideist. Kant did not tolerate theoretical reason in the realm of religion and tried instead to establish his own theology based on practical reason. Therefore, Kantian rational theology is an effort to criticize theoretical reason and confirm moral religion.

# 40. Self-Knowledge and God in the Philosophy of Kant and Wittgenstein

*Chan-Goo Park*

## 1. Relations between the Philosophies of Kant and Wittgenstein

The philosophies of Kant and Wittgenstein seem to be similar in many respects.<sup>1</sup> As such, Geier claims that Wittgenstein is a successor to Kant or even a Kantian.<sup>2</sup> Pitcher also points out that the two philosophies have something in common in their fundamental issues: Wittgenstein “marked out the limits of sense, to indicate the boundary between what can intelligibly be said and what cannot be said. ... In pursuing this task, he was carrying on, in his own way, the work started in modern philosophy by Locke, Hume, and Kant.”<sup>3</sup> In fact, there are documents that indicate Wittgenstein was directly influenced by Kant. For example: “The limit of language is shown by its being impossible to describe the fact which corresponds to (is the translation of) a sentence, without simply repeating the sentence. (This has to do with the Kantian solution of the problem of philosophy.)”<sup>4</sup>

Although Wittgenstein was not well acquainted with the history of philosophy, there is evidence that he had read Augustine, Pascal, and Kant.<sup>5</sup> It is not clear how much of Kant’s writing Wittgenstein had read, but Pitcher guesses Wittgenstein learned Kant’s ideas from his

---

1 Cf. Susanne Fromm, *Wittgensteins Erkenntnisspiele contra Kants Erkenntnislehre* (Freiburg[Breisgau]: Alber, 1979), 1 f.

2 Cf. Manfred Geier, *Kants Welt* (Reinbek bei Hamburg: Rowohlt, 2003), ch. 3.

3 George Pitcher, *The Philosophy of Wittgenstein* (Englewood Cliffs, N. J.: Prentice-Hall, 1964), 326.

4 Wittgenstein, *Culture and Value*, tr. Peter Winch (Oxford: Basil Blackwell, 1980), 10.

5 Cf. S. M. Engel, “Wittgenstein und Kant”, in P. Heintel and L. Nagl (eds.), *Zur Kantforschung der Gegenwart* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1981), 305.

reading of Schopenhauer.<sup>6</sup> Particularly, some philosophers assert that Wittgenstein's conception of the metaphysical subject reflect Schopenhauer's influence, on the basis of the similarity between Wittgenstein's conception of 'I' (*Ich*) and Kant's conception of the "transcendental unity of apperception".<sup>7</sup>

Another view is that we can already find the germ of Wittgenstein's philosophy of language in Kant.<sup>8</sup> Kant's statements below support this view:

[T]he universal and necessary rules of thought in general can concern solely its *form*, and not in any way its *matter*. Accordingly, the science containing these universal and necessary rules is a science of the mere form of our intellectual cognition or of thinking. And we can therefore form for ourselves the idea of the possibility of such a science, just as that of a *general grammar* which contains nothing beyond the mere form of a language in general, without words, which belong to the matter of language.

In addition to the above evidence, those who are interested in the philosophy of Kant and Wittgenstein probably can perceive their similarities in many respects.

#### A. Limitations of the Scientific Intellect

Kant and Wittgenstein had a good understanding of modern science, but also knew its limitations. They seem to have seen the limits of modern science in their own way and warned against the misuse of the scientific intellect. For them, there is a dimension that can be accessible by the scientific intellect and another one that cannot be. One who approaches the latter by a method of the former cannot help making mistakes. We might express the task of Kant's *Critique of Pure Reason* by Wittgenstein's words: in the phenomenal world (i. e., through the scientific intellect) "what can be said at all can be said clearly";<sup>10</sup> but if we

6 Pitcher, 167.

7 Engel, 309.

8 Engel, 315.

9 9:12 f. Immanuel Kant, *Logic*, tr. R. S. Hartman and W. Schwarz (N.Y: Dover, 1974). Quotations of Kant's works in this essay are based on *Werke in zehn Bänden*, ed. W. Weischedel (Darmstadt: Wissenschaftliche Buchgesellschaft, 1983).

10 Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, tr. D. F. Pears and B. F. McGuinness (London: Routledge & Kegan Paul, 1961), Preface.

try to say what cannot be said (through scientific intellect), we inevitably fall in the “fly-bottle” (i. e., the paralogisms or the antinomies).<sup>11</sup>

In addition, both Kant and Wittgenstein approach “what we cannot speak about” in a negative way. “What we cannot speak about we must pass over in silence”,<sup>12</sup> as Wittgenstein stated, can be an example of such an approach. Although this famous statement from the last paragraph of *Tractatus* has been interpreted by logical positivists as “do not discuss metaphysical objects such as ethics, religion, and others that cannot be judged as true or false by means of logic and experience”, it is because they have focused only on the negative meaning and did not recognize the positive meaning of the statement. In other words, they did not comprehend the intention of the saying, to “signify what cannot be said, by presenting clearly what can be said.”<sup>13</sup> We can understand the original intention of the approach from what Kant has stated in the Preface of *Critique of Pure Reason* (Bxxx): “I have therefore found it necessary to deny *knowledge*, in order to make room for *faith*.” We can understand the intention of Wittgenstein in similar terms.

Thus, the aim of this book is to draw a limit to thought, or rather—not to thought, but to the expression of thoughts: for in order to be able to draw a limit to thought, we should have to find both sides of the limit thinkable (i. e., we should have to be able to think what cannot be thought).<sup>14</sup>

Initially, Wittgenstein wanted us to draw the ocean and an island, and as a result, show the shoreline of the island. However, what he actually intended to do was (indirectly) to define the ocean. The fact that he was not satisfied by simply drawing an “island” can be shown in *Tractatus*, as he stated: “We feel that even when all *possible* scientific questions have been answered, the problems of life remain completely untouched.”<sup>15</sup> Similarly, this is shown in *Culture and Value*: “I may find scientific questions interesting, but they never really grip me. ... Fundamentally, I am indifferent to the solution of scientific problems; but not the other sort.”<sup>16</sup>

---

11 Ludwig Wittgenstein, *Philosophical Investigations*, tr. G. E. M. Anscombe (New York: Macmillan, 1953), §309.

12 Wittgenstein, *Tractatus*, 7.

13 Wittgenstein, *Tractatus*, 4.115.

14 Wittgenstein, *Tractatus*, Preface.

15 Wittgenstein, *Tractatus*, 6.52.

16 Wittgenstein, *Culture and Value*, 79.

## B. Transcendental Philosophy and Language Philosophy

One of the main themes of Kant's first *Critique* is to show the limitations and boundaries of the competence of reason. However, Schopenhauer, Kant's successor and critic, changed Kant's question regarding pure reason to a question of concept (*Vorstellung*). In the end, the problem of concept (*Vorstellung*) comes back as the problem of language. As such, the Kantian theme can be said to have been initially revised as 'to show the boundaries and limitations of language' as a result of Schopenhauer's reanalysis. Provided that Wittgenstein was influenced by Schopenhauer as previously mentioned, then it is plausible that he comprehended the Kantian theme as relating to language.

According to Kant, what plays a constitutive role in our cognition is the category of the understanding (*Verstand*) (i. e., *a priori* pure concepts). Similarly, language plays a constitutive role in our thought. In addition, cognition is the task of thought. If so, then it can be said that Wittgenstein has substituted forms of expression (*Ausdrucksformen*) for Kant's forms of judgment (*Urteilsformen*). Wittgenstein's statement regarding this matter, that we need to "draw a limit to thought, or rather—not to thought, but to the expression of thoughts", can be understood in the similar manner. In sum, if Kant intended to show the limitations of reason through his critique of reason and as a result, emphasized its true usage, then we can claim Wittgenstein intended to show the limitations of language and as a result, emphasized its true usage.

Did Kant realize that *a priori* concepts of the understanding can be considered linguistically? As shown in *Prolegomena to Any Future Metaphysics*, Kant was aware of such problem (4:322 f [§39]):<sup>17</sup>

To search in our common knowledge for the concepts which do not rest upon particular experience and yet occur in all knowledge from experience, of which they as it were constitute the mere form of connection, presupposes neither greater reflection nor deeper insight than to detect in a language the rules of the actual use of words generally and thus to collect elements for a grammar (in fact both researches are very nearly related) ...

Let us examine the similarities between these two critical minds. First, these two philosophers, well-known as the sons of modern science, have highly acclaimed Newtonian physics for presenting the phenomenal world as a unified and rational system. Kant's comment that "Newton was the first one to find out the simple orders and regularities of na-

---

17 Cf. Engel, 338.

ture which seems to be disorderly and confusing”<sup>18</sup> and Wittgenstein’s comment that “Newtonian mechanics imposes a unified form on the description of the world”<sup>19</sup> support this claim. However, what these two philosophers focused on was not scientific knowledge itself but the conditions that made it possible for such knowledge to arise. Kant referred to this as transcendental knowledge while Wittgenstein called it grammatical knowledge. The following remarks of Kant and Wittgenstein clearly show the connections between these two types of knowledge:<sup>20</sup>

And here I make a remark, which the reader must bear well in mind ... Not every kind of knowledge *a priori* should be called transcendental, but that only by which we know that—and how—certain representations (intuitions or concepts) can be employed or are possible purely *a priori* (A56/B80).

[O]ur investigation is directed not towards phenomena, but, as one might say, towards the “possibilities” of phenomena. We remind ourselves, that is to say, of the *kind of statement* that we make about phenomena ... Our investigation is therefore a grammatical one. Such an investigation sheds light on our problem by clearing misunderstandings away.<sup>21</sup>

Another similarity between the two philosophers is shown in Kant’s theory of transcendental illusion (*Schein*) and Wittgenstein’s corresponding theory of illusion (*Täuschung*). While the former refers to what occurs when one is outside the conditions that makes knowledge possible (i. e., when one applies any category beyond its proper use in connection with intuition), the latter points out what occurs when one uses words without having a language game where one can properly use the words. In particular, in terms of demonstrating the existence of God, there seems to be general agreement between the mistake Kant has pointed out (i. e., trying to give the objective reality of the phenomenal world to the ideal of pure reason) and Wittgenstein’s understanding of trying to prove the existence of God through factual evidence as the

---

18 “Newton sahe zu allererst Ordnung und Regelmässigkeit mit grosser Einfalt verbunden wo vor ihm Unordnung und schlim gepaarte Mannigfaltigkeit anzutreffen war.” Kant, *Bemerkungen in den “Beobachtungen über das Gefühl des Schönen und Erhabenen”*, ed. M. Rischmüller (Hamburg: Felix Meiner, 1991), 48.

19 Wittgenstein, *Tractatus*, 6.341.

20 Cf. Engel, 307 f.

21 Wittgenstein, *Philosophical Investigations*, §90.

result of confusing the scientific language game with the religious language game.

## 2. The Problem of Self-Knowledge

As we examine the similarities between Kant and Wittgenstein, let us now focus on the problem of self-knowledge. To begin with, let us consider in detail why it is difficult to grasp the concept of “self-knowledge” by means of the scientific or psychological method. The scientific method is based on objective observation. However, in order to observe something, it is assumed that the subject who is observing and the object being observed are separated (that there is spatial distance). The object cannot be seen if it is too closely attached. Accordingly, Kant refers to this space as the “form of outer intuition”. However, I can observe not only the outer world but also myself who is observing. It does not mean I am observing my outer or physical form, but looking inward and observing my inner world (i. e., mind) or my feelings, such as happiness, sadness, hope, despair, and others. This is called “inner intuition”. It allows one to reflect upon one’s consciousness, to be aware of “myself” who is observing something and thinking about something. However, at the moment I see myself, the I who was seen is the I who was thinking just before and not the subject who is thinking at this very moment. This is because the subject who is in activity at the moment is “the I who is seeing” and not “the I who was seen”. In other words, the I who is the object of my inner intuition has become objectified and as a result, has become a past subject. Thus, “I” am not seeing my present self but the past self. As it is important for outer intuition to have spatial distance, inner intuition also needs temporal distance to observe “the self”. Thus, Kant refers to time as the “form of inner intuition”. In this way, science is a knowledge-seeking method that is available within a world that assumes a spatio-temporal form of intuition.

Let us assume I am trying to figure out my mind by using the scientific method. If I attempt to analyze my desire and intention psychologically, just then the “I” as an object of analysis would be the “past I”. It would not be the “present I” who is always capable of making new determinations. Although the “present I” (the conscious self) is presumed to exist, it can never be the object of this type of analysis. If someone tries to regard the self that is analyzed like this as one’s true self, then it could become a way to escape from the self. As a result,



one can easily fall into the trap of self-deception or self-justification. This is because as one says “I know what I am”, it is possible to reject having to confront one’s true self, and as a result, avoid the burdening responsibility of self-determination. In the end, the “self” that is analyzed through scientific method is just the phenomenal self and not the intelligible or noumenal self (B430).

The “new determination” or “self determination” mentioned above is related to free will. The subject of free will is the intelligible self (or the transcendental self). Whereas the empirical self exists within the temporal process or in the phenomenal world, the intelligible subject is the non-temporal self that is not affected by changes in time or physical environment. Therefore, the determination of the free will is an intelligible act that is not an event at a point of time but a form of self awareness of the intelligible self at the very moment. As such, Kant’s way of finding out the “true self” clearly shows the limitations of approaching such a task by means of the scientific method. In other words, it is to profess that “we don’t know nothing but this” or that we cannot understand ourselves beyond this boundary, because in Kant’s Socratic method, recognizing our ignorance is the first step in finding out the truth.

If we summarize what Kant attempted in his three *Critiques* in a phrase, it is that he intended to set aside the contents from our consciousness, that is to empty the mind. It is to set aside empirical contents that fill the conscious activities such as cognitive, moral, and aesthetic judgments. If we remove all of these contents, what is left? Indeed, there is nothing left. This is the reasoning behind the empiricists’ term “tabula rasa”. The contents that will fill the blank slate are one’s own experiences and they will be diverse according to one’s particular time and place. Therefore, the knowledge from such contents can only be relative. However, Kant finds something important in such a blank slate. He refers to it as “form” in contrast to “content”. What he refers to is the competence of the mind itself that can be shown only when the mind is completely emptied of its contents, and he asserts that all human beings possess this general form. If we do not realize that contents based on experience will result only in relativism and skepticism and that universal knowledge cannot be obtained from them, it would be impossible to understand Kant accurately. The profundity of Kantian philosophy is based on this concept of “emptying”. He tried to construct a new met-

aphysics within that emptiness where all of the phenomenal differences and discriminations cease to exist.<sup>22</sup>

We can explain the connotation of Kant's negative approach through his philosophy of subjectivity as follows: the term "Copernican revolution" associated with his philosophy refers to a revolution in our way of thinking. That is, what human beings see and understand are what human beings have made for themselves and exist only to human beings. This might be seen at first sight as an anthropocentric manifestation. However, it actually implies the complete self-realization of human beings regarding the limitations of their consciousness. In other words, human beings should not regard the world they see as something perfectly objective. It also means human beings should not mistake their mathematical or scientific principles for absolute divine truths. That this world is what we have made for ourselves entails on one hand that what is "absolute" cannot be known based on what human beings objectify, and on the other, that this world is nothing but a dream. Accordingly, the "I" in this world is merely the "I" in my dream. The realization that the world and the self I have seen is nothing but a dream allows us to look beyond the limits of the dream. That is because a true reality can never be found in a dream. It can be found only if the consciousness awakens from the dream.

The consciousness that makes dreams in Kantian philosophy is called the transcendental self (i. e., the soul). It cannot be categorized or perceived. It is the basis of a phenomenon but not the phenomenon itself, and it is the basis for a universally valid cognition but cannot be recognized. That is to say, it is the thing-in-itself. This self can be perceived only as a form of self-consciousness like "I am myself", since it is not regulated under specific contents. Indeed, we know a lot about ourselves. However, what we know of ourselves is about the self as an object. In other words, what we know of ourselves is about a phenomenal self, not about the self as a subject that is conscious of the phenomena (i. e., the transcendental self). The self that is conscious of the phenomenal world cannot know about the self itself that observes this world. The eye that looks at the world cannot look at itself and similarly, the self that knows the world cannot know itself.<sup>23</sup>

---

22 Cf. Ja-Koung Han, *An Invitation to Kant's Philosophy* (Seoul: Seokwang-sa, 2006), 6–7.

23 Cf. Han, 272–5.

Wittgenstein was also aware of this problem, as shown in the following:

There is no such thing as the subject that thinks or entertains ideas.

If I wrote a book called *The World as I found it*, I should have to include a report on my body, and should have to say which parts were subordinate to my will, and which were not, etc., this being a method of isolating the subject, or rather of showing that in an important sense there is no subject; for it alone could *not* be mentioned in that book.<sup>24</sup>

The subject does not belong to the world; rather, it is a limit of the world.<sup>25</sup>

Where *in* the world is a metaphysical subject to be found?

You will say that this is exactly like the case of the eye and the visual field. But really you do *not* see the eye. And nothing *in the visual field* allows you to infer that it is seen by an eye.<sup>26</sup>

Wittgenstein's message is clear. He claims that while we need to clarify what can be said through the scientific method, we should also become aware of the limitations of this method. What cannot be explained by means of scientific method should be allowed to make itself manifest. It is to reach emptiness through a constant act of emptying and finally to meet true reality by means of understanding this state of emptiness. However, he is claiming that this cannot be discussed in a positive manner. The following excerpt illustrates how well he understood the instrumental meaning of the knowledge from Zen-Buddhist enlightenment:

My propositions serve as elucidations in the following way: anyone who understands me eventually recognizes them as nonsensical, when he has used them—as steps—to climb up beyond them. (He must, so to speak, throw away the ladder after he has climbed up it.)

He must transcend these propositions, and then he will see the world aright.<sup>27</sup>

We can also hear the Socratic outcry “Know thyself” or ‘Awaken self-ignorance’ in this excerpt. This voice urges us to come into complete self-realization of the emptiness.

---

24 Wittgenstein, *Tractatus*, 5.631.

25 Wittgenstein, *Tractatus*, 5.632.

26 Wittgenstein, *Tractatus*, 5.633.

27 Wittgenstein, *Tractatus*, 6.54.

### 3. Ethics and God

As mentioned above, Kant and Wittgenstein both had profound knowledge of the natural sciences and were deeply impressed by the scientific accomplishments that allowed us to obtain universal and valid knowledge of the phenomenal world. However, according to their critical view, the world we recognize (or describe) by means of the natural sciences arises from the categories (or grammar) provided by the subject; thus, we cannot truly understand the real being of that world, as laid out in its fundamental ground. Kant called this “thing-in-itself” and Wittgenstein referred to it as “what cannot be said” or “the mystical”. In fact, on one hand, this “thing-in-itself” or “the mystical” can be regarded as something that was inevitably assumed in order for the scientific mind to escape from its epistemological dilemma, while on the other, the nature of human beings (or reason), destined to pursue the unlimited, cannot help but deal with it in any way possible. We have discussed this unknowable transcendental self that cannot be recognized; but as we shall see, these two philosophers do not end the story simply by confessing “I do not know myself.”

This search for “the self that cannot be known” or “the true self” becomes an interest for ethics (or values). In particular, it focuses on ethics that has a categorical character. This is because the hypothetical matters that exist in the phenomenal world are reducible to statements of fact. How can we identify an ethics that has this character? Kant expresses this as “a fact of pure reason” (5:47). The fact mentioned here does not refer to what we can experience through our senses but to what we become aware of *a priori* and can inevitably become certain of. However, Kant stated that the justification of the moral law (i. e., the reason we must obey it) cannot be proven. As such, he is acknowledging that a complete deduction of the moral law is not possible. However, in the following passage from *Fundamental Principles of the Metaphysic of Morals*, where this is confessed, we can feel a moral conviction similar to the one from Socratic awakening to our own ignorance. It may be this type of moral conviction that makes Kant’s moral belief possible (4:463): “While we do not comprehend its unconditional necessity, yet we comprehend its *incomprehensibility*, and this is all that can be fairly demanded of a philosophy which strives to carry its principles up to the very limit of human reason.”

Wittgenstein also claims that ethics cannot be put into words and is transcendental.<sup>28</sup> He refers to what cannot be put into words as things that “*make themselves manifest*.”<sup>29</sup> As Kant would say regarding “a fact of practical reason”, Wittgenstein also says the foundation of ethics is provided within us:

If ethics were something that needs to be proven, then, yes! ...

Good and evil appear only through the subject. And that subject does not belong to the world, but it is a limitation of the world.<sup>30</sup>

The subject that recognizes is an empty illusion. However, the subject that wills is given. If there were no will, then there would also be no center of the world, which we call the “I” and is the owner of ethics.<sup>31</sup>

For both Kant and Wittgenstein, therefore, the self (*das Ich*) cannot be recognized but is presupposed and postulated as the foundation for morals. In addition, it manifests itself only through our moral determination.

Where then is the place for God? Kant has declared in the *Dialectic* of the first *Critique* that it is impossible to provide any demonstration for the existence of God from the theoretical perspective. This is because the categories of the understanding that function within our experiential world cannot work for a nonsensible object. However, in the *Dialectic* of the second *Critique* Kant postulates the immortality of the soul and the existence of God through the fact that practical reason necessarily pursues the highest good. As such, it is inevitable for morality to hypothesize the existence of God (5:121–4f). This type of postulation of the existence of God can be seen as a “hypothesis” from the perspective of theoretical reason, but from the perspective of practical reason that imposes a sense of duty to actualize the highest good, it can be called the “faith” of pure reason (5:126).

Kant’s “moral arguments do justify a ‘subjective’ faith, in that they are founded not on objective proof or evidence but on a personal, but rationally commended, decision to adopt a morally upright course of life”<sup>32</sup> For Kant, the soul is immortal and God exists to those who are certain of moral laws. In this perspective, for Kant, God seems to

---

28 Wittgenstein, *Tractatus*, 6.421.

29 Wittgenstein, *Tractatus*, 6.522.

30 Wittgenstein, “Tagebücher 1914–1916”, in *Ludwig Wittgenstein Werkausgabe Band 1* (Frankfurt a. M.: Suhrkamp, 1984), 2. 8. 1916.

31 Wittgenstein, *Tagebücher*, 5. 8. 1916

32 A. W. Wood, *Kant’s Moral Religion* (Ithaca: Cornell, 1970), 34.

exist beyond the phenomenal self but within the non-temporal transcendental self along with moral certainty.

Let us now examine Wittgenstein's interpretation regarding the existence of God. He states in *Culture and Value*:

God's essence is supposed to guarantee his existence—what this really means is that what is here at issue is not the existence of something.

Couldn't one actually say equally well that the essence of color guarantees its existence? As opposed, say, to white elephants. Because all that really means is: I cannot explain what "color" is, what the word "color" means, except with the help of a color sample. So in this case there is no such thing as explaining "what it *would* be like if colors *were* to exist".<sup>33</sup>

Here Wittgenstein is discussing the concept, or grammar, related to the names of the colors or God. We can understand this as follows: those who insist that the color "red" exists and those who insist it does not are both presupposing the conceptual and grammatical existence of color in general or a particular color. Since we are using such words to refer to a specific color, we have within us specific samples of various colors. Indeed, in order for this to be possible, our physiological assumptions should be met (i. e., that we are not color blind).

The problem with the existence of God is similar. The original nature of color or the original nature of God can never be discussed independent from language. We can know what color is and who God is only within a regulated form of language game, or within a common form of life.<sup>34</sup> In fact, in order for communication to be effective, we need to have a universal understanding of a word as well as a common way of using the word. For this commonality to be achieved, we need to live in common grounds, constructing a "shared context". Wittgenstein comments on this notion in the following excerpt:

[T]he *words* you utter or what you think as you utter them are not what matters, so much as the difference they make at various points in your life. How do I know that two people mean the same when each says he believes in God? ... *Practice* gives the words their sense.<sup>35</sup>

For Wittgenstein, faith in God does not add something to life or to the meaning of life but show one's assurance in the meaning of life. The

---

33 Wittgenstein, *Culture and Value*, 82.

34 R. Wimmer, "Gott und der Sinn des Lebens", in *Ankündigung der Sterblichkeit*, ed. J.-P. Wils, (Tübingen: Attempto Verlag, 1992), 105 f.

35 Wittgenstein, *Culture and Value*, 85.

correlation between one's belief in God and one's confidence in the meaning of life is clearly stated in his *Tagebücher*:

To believe in God is to understand the question about the meaning of life.  
To believe in God is to see that it is not yet terminated by the facts of the world.

To believe in God is to see that life has meaning.<sup>36</sup>

A person who lives in harmony with oneself and one's life along with God and the world is a person living in happiness. Such a person no longer doubts the purpose and the meaning of one's existence. Thus, there are no questions to be asked and no answers to be heard. Then, for this person, the phrase "I believe in God" cannot be inferred as meaning "I believe that God exists" or "God exists". It now has to be understood within a religious language game. This type of language game reflects our daily lives and activities.<sup>37</sup>

The transcendental conditions (as Kant puts it) for the form of religious practice cannot be understood by or derived from those that have already been justified. It can only be taken in as it is. According to Wittgenstein, the origins of these various forms of life and forms of practice cannot be regarded as rational or irrational and grounded or not grounded. In this case, a type of firm faith that characterizes a form of religious life is not a rational faith or an irrational faith. We may be able to call it "non-rational" or "trans-rational". Therefore, a form of religious life and practice does not need any justification, nor is it possible for it to have such justification. This is because justification is possible only within an experiential world. The existence of God is no longer an objective matter that is separated from religious life and practice, but a matter that deals with the different types of practice or the various types of meaning in each context of life. Within religious contexts, such a confession that God exists is not a statement formed by a particular concept but is an act of faith. People with such belief are those who are convinced (though there is no reason for supporting it) that their life has an absolute meaning, and that hardship or even death cannot take away such meaning.<sup>38</sup>

---

36 Wittgenstein, *Tagebücher*, 8. 7. 1916.

37 Wimmer, 132 f.

38 Wimmer, 135 f.

#### 4. Concluding Remarks

We have discussed Kant and Wittgenstein's thoughts on the self, ethics and God. Kant examined our competence of cognition which is transcendental in order to make sure that we can have universally valid knowledge. Wittgenstein observed grammar in our language in order to describe the world as it is. However, these two philosophers were much more interested in "denying knowledge" or "drawing a limit to the expression of thought". It was done to "make room for faith" or "signify what cannot be said". In other words, it is because they were interested in enlightening "the noumenon" or "the mystical" that is located on the background of the world that appears through knowledge and language, and that gives meaning to the world.

The key to enlighten the noumenon or the mystical was provided by morality. For Kant, morality is not simply a theoretical but a practical factor that changes the world through practice. In addition, the existence of God is postulated in order to accomplish the commands given out by this moral reason and to see the fruits of such effort. And, God is found to exist already within the transcendental self of those who were certain of moral laws.

To both Kant and Wittgenstein, a person who does one's best morally is the one who is certain of the absolute meaning of life and that this meaning cannot be taken away by hardship or even death. In other words, this is a person who understands the questions regarding the meaning of life and knows that one's own faith cannot be terminated by the facts in this world. Moreover, it is the one who can go beyond the phenomenal self and realize the real self as well as the mystical of the world. Thus, it is a person who has come to an understanding that "It is not *how* things are in the world that is mystical, but *that* it exists."<sup>39</sup>

---

39 Wittgenstein, *Tractatus*, 6.44.



# 41. Kant's Philosophy of Religion as the Basis for Albert Schweitzer's Humanitarian Awareness

Kiyoshi Himi

## 1. The Contents of Schweitzer's Thesis

A philosophy student in Strassburg, Albert Schweitzer chose as the theme of his doctoral thesis Kant's philosophy of religion. His thesis was published by J. C.B. Mohr (Tübingen) in 1899 under the title: *Die Religionsphilosophie Kants von der Kritik der reinen Vernunft bis zur Religion innerhalb der Grenzen der bloßen Vernunft* (*The Religious Philosophy of Kant from the "Critique of Pure Reason" to "Religion within the Boundaries of Mere Reason"*). In this essay I shall argue that his thesis, an excellent study of Kant's philosophy of religion, formed the basis for Schweitzer's own humanitarian awakening a few years later.<sup>1</sup>

The rubrics of Schweitzer's doctoral thesis on Kant are as follows:

Einleitung (Introduction)

1. Teil: Die "religionsphilosophische Skizze" der Kritik der reinen Vernunft (Part 1. The "Sketch for philosophy of religion" in *Critique of Pure Reason*)
2. Teil: Die Kritik der praktischen Vernunft (Part 2. *Critique of Practical Reason*)
3. Teil: Die Religion innerhalb der Grenzen der bloßen Vernunft (Part 3. *Religion within the Boundaries of mere Reason*)
4. Teil: Die Kritik der Urteilskraft (Part 4. *Critique of the Power of Judgment*)
5. Teil: Allgemeine Schlussüberblick (Part 5. Conclusion with a general survey)

Note that the order is reversed between *Critique of the Power of Judgment* (CJ) and *Religion within the Boundaries of mere Reason* (RBR). According to the chronological order, Part 3 should treat of the former, published in 1790, and then Part 4 should treat of the latter, published three years later. Seemingly, Schweitzer defies the chronological order. That might perplex us all the more because he entitles the book: *The Religious Phi-*

---

1 As to Schweitzer's biography, especially his medical activities, cf. *Albert Schweitzer: An Anthology*, ed. Charles R. Joy (U.S.A.: The Beacon Press, 1947), 291–307.

*osophy of Kant from the "Critique of Pure Reason" to "Religion within the Boundaries of Mere Reason". How does he dare to contradict himself?*

As we digest Schweitzer's text, we can understand his intention: he purposefully reverses the chronological order to contrast *RBR* with *Critique of Practical Reason (CPrR)*. As we know, not a few people who study Kant's philosophy of religion focus their mind on *CPrR*, because they suppose Kant included in that work his complete teachings on religious philosophy, as he had mapped out in *Critique of Pure Reason (CPR)*. Yet Schweitzer takes a different view. In his estimation *CPrR* faultily constructed Kant's philosophy of religion, because he could there only partially take on his plan in *CPR*. By contrast, in *RBR* Kant succeeded in fully developing that plan. So Schweitzer highlights the distinction between both works. After that he describes *CJ* as the intermediary work where Kant could revise his view and prepare for his argumentation in *RBR*.

Schweitzer considers Kant's concept of freedom as a clue to the study of Kant's philosophy of religion. So he analyzes in the first part of his thesis Kant's description of the concept of freedom in *CPR*. Then, in the second part he argues that Kant's argumentation in *CPrR* only partially embodies his concept of freedom as planned in the previous work. In the third part he contrasts *RBR* with *CPrR* and shows how Kant succeeded this time in embodying his first planned concept of freedom. In the fourth part he argues that Kant's teleological consideration in *CJ* is what enabled him to proceed with improving his grasp of the concept of freedom.

## 2. The "Sketch for Philosophy of Religion" in *Critique of Pure Reason*

Schweitzer characterizes the chapter "The Canon of Pure Reason", the third last chapter of *CPR*, as "die 'religionsphilosophische Skizze'", that is to say, the "Sketch for philosophy of religion". As Schweitzer remarks, Kant mentions there three objects that our reason in its transcendental use is directed to in the end: the freedom of the will, the immortality of the soul, and the existence of God. However, Kant is at the same time emphatic that reason in the merely speculative regard has little interest in these objects, because it has in fact no hope of making any use of the discoveries about them, if such are attainable at all. So the im-

portance of these cardinal objects to our reason does not ultimately concern knowing, but has another, practical purpose. In this view Kant explicates the concept of freedom, regarded as the most fundamental of the three, emphatically in its practical sense, differentiating this from its transcendental sense.

Supposedly, that chapter derives from an earlier draft written in the 1770s. Of course, Kant supplemented it to make it cohere with the context of the concluding part of *CPR*. Consequently, Kant's "Sketch for philosophy of religion" contains two incompatible factors: (1) the remainder of his earlier thoughts of the 1770s and (2) the argument that relates to the theory he elaborated anew in the preceding parts of *CPR*. His argument about the concept of freedom directly reflects this background. In fact, "the concept of freedom in its practical sense", or "practical freedom", represents the former factor, and "the concept of freedom in its transcendental sense", or "transcendental freedom", represents the latter factor (A803–4/B831–2). What is most interesting is that Kant asserts here that transcendental freedom, whose possibility he showed in *Transcendental Dialectic*, is irrelevant to the inquiry into practical freedom. According to him, reason can give the imperatives and precepts for conduct based on its peculiar consideration and thus determine human choice. In our daily life we are conscious of that causality of reason. In that sense practical freedom can be proved through experience. Whether in these actions reason is itself determined by further influences, or whether it is absolutely free in initiating a series of events, does not concern us insofar as we would ask a practical question.<sup>2</sup>

---

2 Kant tells as follows A802–3/B830–1: "Practical freedom can be proved through experience. For it is not merely that which stimulates the senses, i. e., immediately affects them, that determines human choice, but we have a capacity to overcome impressions on our sensory faculty of desire by representations of that which is useful or injurious even in a more remote way, but these considerations about that which in regard to our whole condition is desirable, i. e., good and useful, depend on reason. Hence this also yields laws that are imperatives, i. e., objective **laws of freedom**, and that say **what ought to happen**, even though perhaps never does happen, and that are thereby distinguished from **laws of nature**, which deal only with that **which does happen**, on which account the former are also called practical laws. But whether in these actions, through which it prescribes laws, reason is not itself determined by further influences, and whether that which with respect to sensory impulses is called freedom might not in turn with regard to higher and more remote efficient causes be nature—in the practical sphere this does not concern us, since in

Obviously, Kant has yet to explain the concept of the moral law as the autonomy of the will. As he concedes, practical freedom cannot prove the faculty of reason to initiate a series of appearances, while transcendental freedom requires just that, insofar as its possibility was acknowledged through the solution of the third antinomy. However, Kant asserts there is no need for practical freedom to be adequate for the transcendental signification of the concept. According to him, it makes no difference whether reason is absolutely spontaneous, or is itself determined by further influences. He is convinced we can dispense with transcendental freedom in the field of practical philosophy. Schweitzer lays emphasis upon this original stance of Kant's.

### 3. Transcendental Freedom and Practical Philosophy

Had Kant proceeded with his practical philosophy according to the "sketch", his second *Critique* would not have appeared. He could have written moral philosophy and explicated practical freedom without relating it to transcendental freedom. Yet, in actuality, he discussed the problem of practical freedom in *CPrR*. This indicates that he had to connect practical freedom, against his own expectations, with transcendental freedom and to construct his practical philosophy on that basis. In other words, he imposed restrictions upon practical freedom in order to put it into the framework of his transcendental idealism. How did that come about?

As is well-known, it was in the Transcendental Dialectic that Kant proposed transcendental freedom as a possible concept. He elaborated the concept through the solution of the third antinomy of pure reason.<sup>3</sup> In the antinomy the thesis and the antithesis try to refute each other, but cannot provide any conclusive proof. So the dispute seems to be endlessly undecided. This situation is common to all the four types of antinomy disclosed in *CPR*. Yet the solution Kant gives to the third an-

---

the first instance we ask of reason only a **precept** for conduct; it is rather a merely speculative question, which we can set aside as long as our aim is directed to action or omission."

3 It runs as follows (A444-5/B472-3): "Thesis: Causality in accordance with laws of nature is not the only one from which all the appearances of the world can be derived. It is also necessary to assume another causality through freedom in order to explain them." "Antithesis: There is no freedom, but everything in the world happens solely in accordance with laws of nature."

tinomy is unique: he admits both the thesis and the antithesis to be true and mediates the case to the satisfaction of both parties. The clue to that solution is that in the causal connection the condition is not necessarily homogeneous with the conditioned, but can also be heterogeneous with it—that is, not of the same kind: every occurrence in the world of appearance is determined by another preceding appearance; this is again determined by a further preceding appearance, and so on. In our exploration of nature we must regard this series of temporal conditions (i. e., the causal connection of appearances) as uninterrupted, thoroughgoing. So what the antithesis contends is justified. Yet, at the same time, an occurrence can have a heterogeneous cause that may be beyond the world of appearances—that is, it may belong to the intelligible world. By ascribing freedom to this cause we can justify the contention of the thesis.<sup>4</sup>

Insofar as “the antinomy of pure reason” is presented as the critique of “rational cosmology” (A408–20/B435–48), the dispute should be related to the whole world. Kant also examines in the third antinomy the causal connection of things in general. But it is quite obvious that his concern is just one sphere of the universe: human action. The double aspect of causal determination, whereby Kant tries to save freedom, can apply only to the sphere of occurrences called human actions. A human action as an occurrence in the world is determined by a preceding appearance and thus belongs to the causal connection of all the appearances in accordance with natural laws. At the same time, a human action can be regarded as being determined by a “heterogeneous”

---

4 Kant explains that as follows (A538–9/B566–7): “I call **intelligible** that in an object of sense which is not itself appearance. Accordingly, if that which must be regarded as appearance in the world of sense has in itself a faculty which is not an object of intuition through which it can be the cause of appearances, then one can consider the **causality** of this being in two aspects, as **intelligible** in its **action** as a thing in itself, and as **sensible** in the **effects** of that action as an appearance in the world of sense. Of the faculty of such a subject we would accordingly form an empirical and at the same time an intellectual concept of its causality, both of which apply to one and the same effect. Thinking of the faculty of an object of sense in this double aspect does not contradict any of the concepts we have to form of appearances and of a possible experience. For since these appearances, because they are not things in themselves, must be grounded in a transcendental object determining them as mere representations, nothing hinders us from ascribing to this transcendental object, apart from the property through which it appears, also another **causality** that is not appearance, even though its **effect** is encountered in appearance.”

cause—that is, a motive of the will, independent of all the causality of nature. Kant suggests that this motive of the will originates in the intelligible character of the human being—that is, in the human being as “thing in itself”—and characterizes its causality as transcendental freedom.

Transcendental freedom, as the faculty to initiate a series of appearances or to originate a causal chain of events, is therefore proved to be possible in the sphere of human actions. Yet it is still a probable concept. In the next stage its reality must be proved. This task must be entrusted to the practical use of reason. So, in spite of Kant's statement in *The Canon of Pure Reason*, his practical philosophy can by no means dispense with transcendental freedom, insofar as he would keep the results obtained through the solution of the third antinomy. In this view mere practical freedom, as was described in that chapter, is insufficient and falls short of the philosophical conception of freedom. On the contrary, Kant is now faced with the problem of proving the reality of transcendental freedom.

Proceeding with the considerations of practical philosophy, Kant occupied himself in distinguishing between causes that derive from reason. If he would succeed in sifting out the causes that signify the absolute spontaneity of reason from other causes, the former would be able to substantiate transcendental freedom. In fact, in *Groundwork of the Metaphysics of Morals (GMM)* Kant drew an important distinction between the hypothetical and the categorical imperative. As long as the reason causing an action is determined by another cause, its lawgiving function is conditioned and the imperative it prescribes is hypothetical. Contrariwise, if reason causes an action quite independently of other influences and thus initiates a series of causes, the imperative it prescribes is categorical. In the latter case, reason gives arbitrarily and spontaneously a law for action and initiates a series of causes. So we acknowledge the act of reason as autonomy, whereas in the former case we justly call it heteronomy. Thus, the categorical imperative indicates freedom of reason, while in the hypothetical imperative reason serves another cause as a means. Kant contends that the only genuine moral imperative must be categorical. It originates in the autonomy of reason and meets the requirement of transcendental freedom.

#### 4. Philosophy of Religion in *Critique of Practical Reason*

Kant was convinced he succeeded in *GMM* in connecting freedom with morality and thus getting ready for demonstrating transcendental freedom in its reality. He founded his principles of practical philosophy in *CPrR* on this base. At the beginning of the book he describes freedom as inseparable from the moral law. He contends that we can become aware of freedom only on the basis of our consciousness of the moral law.<sup>5</sup> The moral law is the rationale for our awareness of freedom (*ratio cognoscendi*), while freedom founds the existence of the moral law (*ratio essendi*). Thus he explicates the interdependence between freedom and the moral law. The moral law manifests itself as the categorical imperative, originating in the autonomy of reason. Kant describes it as “the sole fact of pure reason” (“das einzige Faktum der reinen Vernunft” [5:31])—the only fact that announces itself commonly to all human rational beings. On this “popular” fact he grounds his arguments in practical philosophy.

The reality of freedom is thus proved by the fact of the moral law and the human will is recognized as a free will, insofar as it accepts the moral law as the motive for action. Kant then proposes as the object of that free will the idea of “the highest good”: the complete harmony of one’s disposition with the moral law, and the proper happiness distributed to each person in accordance with his/her worthiness. With regard to the former we must suppose our soul will exist beyond our bodily death and continue making efforts to improve our disposition, and with regard to the latter we must suppose the existence of an omniscient

---

5 He contends as follows (5:4): “freedom is real, for this idea reveals itself through the moral law”; and “among all the ideas of speculative reason freedom is also the only one the possibility of which we *know* a priori, though without having insight into it, because it is the condition of the moral law, which we do know.” The following is also an important remark (5:4): “Lest anyone suppose that he finds an *inconsistency* when I now call freedom the condition of the moral law and afterwards, in the treatise, maintain that the moral law is the condition under which we can first *become aware* of freedom, I want only to remark that whereas freedom is indeed the *ratio essendi* of the moral law, the moral law is the *ratio cognoscendi* of freedom. For, had not the moral law *already* been distinctly thought in our reason, we should never consider ourselves justified in *assuming* such a thing as freedom (even though it is not self-contradictory). But were there no freedom, the moral law would not *be encountered* at all in ourselves.”

and omnipotent world-ruler who distributes the proper happiness to each person. In other words, we postulate the immortality of our soul and the existence of God. Kant defines a postulate as a “*theoretical proposition*, though one not demonstrable as such, insofar as it is attached inseparably to an *a priori* unconditionally valid *practical law*” (5:122). Since the highest good is the ultimate object of the will determined by the moral law, its constituents, the immortality of the soul and the existence of God, are the objects of postulation. Moreover, in the Dialectic of Pure Practical Reason Kant counts one more subject among the postulates: freedom considered positively. Having already mentioned in *CPR* freedom of the will, the existence of God, and the immortality of the soul (or a future life) as three ultimate concerns of pure reason, he solved that problem in *CPrR* in the form of the system of postulates that can also be characterized as the creeds of moral religion. Kant seems to have believed at that time that he was successful in this solution. Indeed, not a few people regard *CPrR* as the climax of his philosophical development.

Yet Schweitzer refutes the overestimation of *CPrR*, pointing out the faults of Kant's argument. First, from the standpoint of logical analysis Kant commits an obvious error. As remarked above, he counts freedom among the postulates. However, that betrays his self-contradiction, for the reality of freedom is already confirmed in the Preface of the book as the ground for the existence of the moral law as such. On the base of freedom the necessity of the highest good is recognized and its constituents, the immortality of the soul and the existence of God, can be postulated. Therefore freedom cannot properly belong to the same class as the other two, nor can the name of postulate fit it. In fact, Kant only patches up his argumentation in order to give an appearance of an orderly system of postulates to his religious philosophical worldview.<sup>6</sup>

Secondly, what is more important, Schweitzer's critique pertains to the isolation of the individual. According to Kant, freedom as the faculty of reason to initiate a series of appearances—that is, as the answer to the thesis of the third antinomy—is cognizable through the moral law, and every human individual acknowledges inside of him/herself its motivation to act. Therefore, Kant asserted that an individual acts morally and from freedom, insofar as his/her will complies with the moral law and thus constructs the maxim. But practically, that just de-

---

6 Cf. Schweitzer, *Die Religionsphilosophie Kants*, 80–2.



pende on his/her own persuasion. The same action the individual asserts to be motivated by the moral law, and therefore free, is as an occurrence in the sensible world determined by a preceding appearance and thus belongs to the causal connection of all the appearances in accordance with natural laws. Others cannot but estimate it as an appearance in the causal chain of events. In this way moral freedom is just a matter of compliance under the condition of the identity of the acting and judging individual. An individual, insofar as he/she is a “moral” actor, is necessarily isolated from others in asserting his/her action to be motivated by the moral law. Moreover, the individual is regarded in the moment of action as being conditioned by two worlds: insofar as his/her will complies with the moral law, he/she acts as “thing in itself” and therefore belongs to the intelligible world, whereas, insofar as the action is determined through the causality of nature, he/she is an appearance, a phenomenon, and therefore belongs to the sensible world. Schweitzer points out that these difficulties necessarily arose because Kant forced the solution of the third antinomy, the double aspect of causal determination, onto the moral law.<sup>7</sup> The framework of transcendental idealism limited moral freedom to the perspective of an isolated individual, split into two worlds.

So in Schweitzer’s view it was necessary for Kant to rectify his argumentation in *CPrR*, as he afterward admitted its fault. Then he could proceed again with developing his argumentation according to his “sketch” in *CPR*. Schweitzer insists that Kant finally succeeded in proposing his sufficient concept of freedom, and the religious world-view based on it, in *RBR*, especially in its third part.

## 5. The Concept of Freedom and the Idea of Moral Community in *Religion*

It is noteworthy that in *RBR* Kant relates human freedom not to the will (*Wille*), but to the power of choice (*Willkür*) and points out that the human being’s power of choice has a propensity to prefer the sensory motives to the motive of reason, that is, the moral law. Since in his view this propensity indwells in human nature, he characterizes it as the

---

<sup>7</sup> Cf. Schweitzer, *Die Religionsphilosophie Kants*, 82–6.

radical evil. His central consideration is how we can surmount it. Schweitzer appreciates this book just in this respect.

In the "sketch for philosophy of religion" Kant already based practical freedom upon the motivation of reason. He stated that, among the various motives whose presence we become conscious of when we are going to act, there are certainly some that can only be represented by reason. Therefore, the causality of reason can be cognized through experience and it grounds practical freedom. So, according to the plan, he should have examined in his practical philosophy the human being's power of choice, in order to prove the reality of practical freedom. He should have shown how our power of choice can choose the motive of reason, rejecting every sensory impulse and letting the causality of reason come into play. Yet in *CPrR* he ascribed freedom to the moral will, as equated with the lawgiving reason. This was caused by the adaptation of the moral law for transcendental idealism: as he found a pure motive of reason in the moral law, he combined it with transcendental freedom, providing the full solution of the third antinomy. As a consequence, the moral law can be regarded as testifying to reason's faculty to initiate a series of appearances. Insofar as the will complies with the moral law, it is admitted to be the free cause of action coming from the actor's intelligible "thing in itself".

After this deviation in *CPrR*, Kant at last makes an issue of the power of choice in *RBR*. He now explicates human freedom as the choice between motives for the construction of the maxim. One does good from freedom, if one accepts the moral law into the maxim as the primary motive, but does evil from freedom, if one lets other motives take precedence over the moral law. Kant recognizes that the human being has an inveterate habit to prefer other motives to the moral law. He regards this as a propensity to evil (i. e., radical evil) in human nature. Now his concern is how the human being can overcome radical evil and restore the predisposition to good to its power. Of course, he admits that one can only little by little improve in one's disposition and that there is no knowing when one can root out the propensity to evil. So it seems that moral progress also requires here an endless time, just as in *CPrR*. However, this requirement does not lead this time to the postulate of the immortality of the individual soul. Kant asserts that God, whose intellectual intuition in an instant grasps the temporal process as a whole, justifies the person who makes every effort to improve his/her disposition, although in his/her own eyes he/she is only on the way of a gradual improvement. Moreover, Kant declares

that moral completion through the defeat of evil is the duty of humankind. An individual may die without attaining to perfection, but he/she can participate in his/her way in the efforts of the whole of humankind to fulfill its duty—that is, to accomplish its objective of moral completion. In this sense the individual can become “immortal”.

In Schweitzer’s estimation Kant in *RBR* thus bases the completion of morality, or the highest good, upon the relationship between God and humankind. Kant prepared the way for this basis in *CJ*: in the second part of that book, having confirmed the validity of teleological judgment in the cognition of organisms, he explicated teleology as a whole. He recognized that teleology in its full development constructs a systematic worldview; at the top stands the being with the highest value in this world: the human being under moral laws. Kant characterized this as “the final end of creation” (5:421–2). But, as Schweitzer emphasizes it, we must interpret here “the human being” not as an individual, but as a collective concept. We as humankind should remodel the world in order that all things should be properly subordinated to our moral purpose. This activity creates culture out of nature and will bring the teleological world to completion. In *CJ* Kant also proposed a proof of the existence of God in the form of the inference from the “final end of creation” to the Creator. He named it “the moral proof of the existence of God”. God, whose existence is now proved as the creator of the moral world, rules the world under moral laws and keeps it in the moral order. Theology, as founded on this proof, is called “ethicotheology”. We can recognize with Schweitzer that the ethico-theological worldview forms the background of the argumentation in *RBR*.<sup>8</sup>

Kant did not plan ahead to publish *RBR* in book form, but intended to have four articles separately published in *Berliner Monatsschrift* (*Berlin Monthly Journal*). Nevertheless, as Schweitzer’s argument in Part 3 shows it, he regards the book as thoroughly consistent in its theme.<sup>9</sup> It follows an orderly sequence: (1) the affirmation of the human being’s ability to improve in the disposition and to attain to moral perfection

8 Schweitzer, *Die Religionsphilosophie Kants*, 290–311.

9 The rubrics of the book are as follows (6:15): “Part 1: Concerning the indwelling of the evil principle alongside the good, or, Of radical evil in human nature.” “Part 2: Concerning the struggle of the good with the evil principle for dominion over the human being.” “Part 3: Concerning the victory of the good over the evil principle and the founding of a Kingdom of God on earth.” “Part 4: Concerning service and counterfeit service under the dominion of the good principle, or, Of Religion and Priestcraft.”

in spite of the persistence of radical evil, (2) the manifestation of an example of a perfectly moral person, so that human beings can expect to overcome the evil, (3) the foundation of a moral community as a people of God under moral laws by human beings imitating the example, and (4) the warning against degradation of the well-founded community. Since the last part can be regarded as additional, we can regard the third part as the climax of the development. I think this is just the point Schweitzer focuses on in Part 3 of his book.

In the first part of *RBR*, after having described the obstinacy of the propensity to evil, Kant shows the possibility for the human being to overcome it and to restore the predisposition to good to its power. Then, from the second part on, he explains how human beings can defeat the dominion of the evil principle and bring about the victory of the good principle. In the second part he speaks of “the personified idea of the good principle” (“personifizierte Idee des guten Prinzips” [6:60]) that should function as an example of complete moral personhood to all human beings. Obviously, Kant means by that the personhood of Jesus Christ, although he never mentions the name. In the third part he at last explicates the foundation of the moral community of humankind. Since the example is already manifested, human beings should imitate it and endeavor to improve in their own disposition, serving each other in every way morally. Their cooperation forms a universal community that Kant calls a people of God under moral laws. Since human beings are the final end of creation only as humankind under moral laws, the foundation of this community is nothing other than the performance of the mission that God has entrusted to them. Schweitzer argues that in this argument Kant's former individualistic tendency, conditioned by transcendental idealism, is cleared away and replaced by the idea of solidarity.

## 6. Conclusion

Although Schweitzer did not write any monographs on Kant's philosophy other than his doctor thesis, his contribution to the study of this subject is indisputable. We acknowledge his argument about Kant's deviation from the “sketch for philosophy of religion” in *CPrR* and then recovery in *RBR* to be cogent. Moreover, it is noteworthy that he drew from his study of Kant's philosophy the idea of a moral community of humankind. He recognized with Kant the foundation of this commun-

ity as the duty of humankind as “the final end of creation”. That idea depicted for him precisely the world into which he would launch out some fourteen years later.

In conclusion I would like to raise one more issue: Christianity forms the background of Kant’s philosophy of religion and Schweitzer adopted Kant’s position. In the first part of *RBR* Kant compares radical evil, for the convenience of explanation, to the original sin. The argument about “the personified idea of the good principle” in the second part of the book implies, as remarked above, the personhood of Jesus Christ deprived of the mythology. And in the third part Kant models the founding and the development of the moral community of humankind upon the Church and its history. Although Schweitzer insisted in his thesis that we should draw a distinction between Kant’s philosophy of religion and Christian dogma, yet he was far from refusing Christianity as the background of Kant’s as well as of his own thought. It was on the basis of Christianity that he adopted from Kant the religious world-view for humankind. So the moral community of humankind meant for him in the last analysis an expansion of the Christian brotherhood. Later in his medical service this pattern for thinking also determined his attitude toward the native Africans. He said he regarded them as children with whom “nothing can be done without the use of authority”, and with regard to them he coined the formula: “I am your brother, it is true, but your elder brother.”<sup>10</sup> Certainly, we can concede that he uttered with that statement an attitude of mind necessitated by the situation he was put in at that time. However, we must also remark that his attitude was still restricted by the Western ideology of Christian superiority and he could not be an egalitarian in the full meaning of the word.

---

10 Cf. Albert Schweitzer, *Philosopher, Theologian, Musician, Doctor—Some Biographical Notes*, compiled by C. T. Campion (London: A. & C. Black Ltd., 1928), 24–5.

## 42. Kant's Religious Perspective on the Human Person

*Brandon Love*

### 1. Introduction

In a recent book entitled *Kant and Theology at the Boundaries of Reason*,<sup>1</sup> Chris Firestone makes the connection between Stephen Palmquist's interpretation of Kant's transcendental philosophy and the theology of Paul Tillich. This essay explores this connection in relation to the issue of human religious identity. My goal in this essay is to illustrate Kant's religious perspective on the human person by using the elements of Kantian rational religious faith as a corrective to the theological system of Paul Tillich. Toward the end of his life, Tillich, as a Christian theologian and Kantian, made a turn toward a naïve religious pluralism. It is my contention that this maneuver is unwarranted given his Kantian foundation. I argue that, in order to be more authentically Kantian, Tillich's theology of religions should adhere to the elements of Kant's rational religious faith. While acknowledging the many insightful elements in Tillich's theological system, I examine his theology in relation to Kant's philosophical program, especially regarding the issue of human personhood. Then, I offer a rejoinder to the religious pluralism of Tillich's system in the form of Kant's rational religious faith.

In order to make my case, I first give an overview of Kant's philosophy, viewing the entire Kantian philosophical enterprise as groping toward the answer to Kant's question regarding the nature and destiny of humankind. In doing so, I outline a theologically affirmative<sup>2</sup> method of Kant interpretation, using Palmquist as the main rudder, and show im-

---

1 Chris Firestone, *Kant and Theology at the Boundaries of Reason* (Aldershot, U.K.: Ashgate Publishing Ltd., 2009).

2 For a discussion of the theologically affirmative interpretation of Kant, see Chris L. Firestone and Stephen R. Palmquist, "Editors Introduction" in Chris L. Firestone and Stephen Palmquist, *Kant and the New Philosophy of Religion* (Bloomington: Indiana University Press, 2006).

portant themes that tie Palmquist and Tillich together. With this method in hand, I examine the relevance of Tillich as a theologian applying Kant in every step of his theological enterprise, especially pertaining to the issue of human religious identity. This paves the way for a comparison of Tillich's theology with Kant's rational religious faith. A careful examination of these two thinkers, I contend, offers a fruitful and creative alternative to Tillich's pluralistic move.

In short, we will see that Tillich's move toward pluralism is unwarranted by his Kantian foundation. Rather, what we will find is that Kant offers us a critical alternative to both full-blown religious pluralism and a more traditional religious exclusivism. My contention is that, rather than follow Tillich toward a full-blown pluralism, Kant, interpreted in a broadly Palmquistian style, offers a viable and, until now, untried alternative. Kant terms it rational religious faith. Though the exact nature of Kantian rational religious faith is still under dispute in the literature, its contours are becoming clear, as I will illustrate. As such, it defines, or at least should define, human self-identity at the transcendental boundaries of reason.

## 2. The Palmquistian Interpretation: Kant's Program as the Search for Human Nature

In this section, I will give a brief outline of Kant's philosophy as interpreted in a theologically affirmative manner in the sense of Palmquist. In *Kant's System of Perspectives*,<sup>3</sup> Palmquist sees Kant's philosophical program as having three main levels. The first level is the overall Transcendental or Copernican Perspective; the second is that of the three *Critiques* comprising, for Palmquist, the distinct standpoints of the system; and the third is the innermost level of the perspectives that arise and are at work in each of the three standpoints. Although the level of perspectives obviously is crucial to understanding Palmquist's system of perspectives in its entirety, here I will focus only on the first two levels, as the overall picture of the system as it relates to Tillich is all that is relevant to the present undertaking.

The overarching Copernican/Transcendental Perspective of the system is based on the turn Kant takes in relation to the history of phi-

---

3 Stephen R. Palmquist, *Kant's System of Perspectives: An Architectonic Interpretation of the Critical Philosophy* (Lanham, Md.: University Press of America 1993).

losophy before him. Kant called this turn the Copernican Revolution in philosophy. This is why Palmquist calls this perspective the Copernican Perspective; he also refers to it as the Transcendental Perspective because it "is based on the assumption that the subject imposes certain a priori forms on the object."<sup>4</sup> This is the basic turn to the subject, hence the Copernican reference. The starting point for Kant's philosophy is the subject's imposition on the object. In the Palmquistian architectonic, this turn "defines the systematic context into which all three Critical systems fit."<sup>5</sup> The second layer of the system is that of the three *Critiques* themselves.

Kant's program is centered on four distinct but inseparably related questions: (1) What can I know? (2) What ought I to do? (3) What may I hope? and (4) What is man? Kant's goal is to answer the fourth question in light of the first three. This is why Kant's philosophy can be seen as a striving after the nature of humanity. Also, it is from these questions that the three basic metaphysical ideas of God, freedom, and immortality arise. These questions and the ideas of reason associated with them are the keys to understanding the standpoints of Palmquist's system. The first question, "What can I know?", corresponds to the theoretical standpoint of the first *Critique*. It is here that Kant defines the boundaries of human knowledge. The second question, "What ought I to do?", corresponds to the practical standpoint of the second *Critique*, where Kant works within the limits of the theoretical standpoint while focusing not on knowledge but on moral action. The judicial standpoint of the third *Critique* and *Religion* is an attempt to bridge the gap of the divide between facts and values arising in light of the answers given in the first two *Critiques*. This is an attempt to answer the question, "What may I hope?", with regard to my moral striving in the face of a world seemingly devoid of ultimate meaning.

Kant's three questions jointly lead to the ultimate question, "What is man?" This question corresponds to the overarching Copernican or Transcendental Perspective. There is no specific *Critique* for this Perspective; rather all three *Critiques* or standpoints, taken together, make up this Perspective. In this way, the *Critiques* have a twofold function in Kant's philosophy. Each *Critique*, and corresponding standpoint, is a system on its own and also a part of the overall Copernican/Transcendental System. The goal of each individual standpoint is to answer its

---

4 Palmquist, *Kant's System of Perspectives*, 58.

5 Palmquist, *Kant's System of Perspectives*, 58.



particular question (whether of knowledge, morality, or hope), while at the same time to contribute to the overall answer of the entire system (“What is man?”) and its turn to the subject.

In this quest for hope and meaning, the practical standpoint, in relation to the judicial, has primacy over the theoretical. This is due to the fact that Kant is seeking hope in light of our moral striving, and meaning, for Kant, is to be found in light of values rather than facts. Kant finds that the world of facts does not line up with that of values; in short, the world is not the way it ought to be. This is the reason Kant begins his quest for hope. Kant turns to the notion of the highest good.<sup>6</sup> This leads to the postulation of God as lawgiver and judge. However, Kant, in *Religion*, has come to the realization that the issue of hope is still unsettled. Here he turns to the issue of religion and lays out his notion of rational religion. We will examine this notion more closely in the rejoinder to Tillich. Basically, Kant claims that humanity is evil and in need of grace; he then postulates the idea of the prototype of perfect humanity and claims that humanity, if we are to have hope, must undergo a conversion by conforming to the prototype; lastly, Kant says that we must band together into ethical communities or “churches” in order to combat evil. We will see how these elements can be applied to a philosophical conception of religion in the rejoinder to Tillich; however, an outline of the relevant aspects of Tillich’s thought must first be presented. To this task I now turn.

### 3. Tillich on Human Personhood and Revelation: The Unfounded Pluralist Foundation

It is well-known in theological circles that Tillich was a theologian writing in the wake of Kant. Tillich’s conception of his task as a theologian is based upon Kant’s distinction between the philosophy and theology faculties in his *Conflict of the Faculties*; also, Tillich’s notion of the theological circle is a reworking of Kant’s vision of philosophy and theology as comprising concentric circles, presented in *Religion within the Bounds of Mere Reason* (or *Religion*). An as yet unexplored aspect of this relationship is Kant’s philosophical doctrine of rational religious faith in com-

---

6 See Frederick C. Beiser, “Moral Faith and the Highest Good”, in Paul Guyer (ed.), *The Cambridge Companion to Kant and Modern Philosophy* (Cambridge: Cambridge University Press, 2006).

parison with Tillich's theological system, especially Tillich's view of the doctrine of revelation. As we have seen, Kant, in *Religion*, lays out some doctrines required of any religion in order for it to be considered rational. Toward the end of his life, Tillich flirted with Eastern religions. Today, a growing movement in Tillich-studies believes this period of his work lays a foundation for religious pluralism. My contention is that Tillich, in this period, gave in to a naïve philosophical conception of religious pluralism that lacked Kant's criteria for rational religion.

Just as practical reason is primary in Kant's philosophy, the cornerstone of Tillich's theory of human nature is his ethical thought, especially the moral imperative. Tillich says "the moral imperative is the command to become what one potentially is, a *person* within a community of persons."<sup>7</sup> He goes on to say of this moral imperative toward the realization of personhood:

The moral imperative is the demand to become actually what one is essentially and therefore potentially. It is the power of man's being, given to him by nature, which he shall actualize in time and space. His true being shall become his actual being—this is the moral imperative. And since his true being is the being of a person in a community of persons, the moral imperative has this content: to become a person. Every moral act is an act in which an individual self establishes itself as a person.<sup>8</sup>

Tillich speaks of the notion of a person as a person, the choices or moral actions of the individual self as establishing oneself as a person, and the fact that the person is necessarily a person in a community of persons. This is a standard statement from Tillich concerning personhood. It is mainly concerned with the existential and communal elements of ethics. However, while Tillich's thought on the ethical, existential nature of human persons is foundational, he spends the majority of his time and energy describing the ontological nature of humanity as religious.

This other, fuller and more developed, aspect of Tillich's notion of personhood deals with the ontological nature of humanity. For Tillich, this is mankind's relation to the mystery of Being or Being-itself. In Tillich's thought,<sup>9</sup> God is Being-itself and the ground of Being, while humanity, as estranged from our essential Being, is able to participate

---

7 Paul Tillich, *Morality and Beyond* (Louisville: Westminster John Knox Press, 1995), 19.

8 Tillich, *Morality and Beyond*, 20.

9 Paul Tillich, *Systematic Theology*, 3 vols. (Chicago: University of Chicago Press, 1951, 1957, and 1963, respectively).

in Being. This participation has many facets; not least of these ways is humanity's participation in Being-itself through revelation (the revelation of Being-itself to and in humanity). For Tillich, revelation is an aspect and natural (or supernatural) product of reason. He divides reason into two basic, distinct-but-related parts. The first is technical reason. This is objective reason or simply the capacity for reasoning (i.e., logic). The other sense of reason, ontological reason, is subjective. It is also a fuller concept of reason for Tillich. He claims that ontological reason is "the structure of the mind which enables the mind to grasp and to transform reality."<sup>10</sup> This springs from Kant's transcendental philosophy; Tillich views it as the existential process that includes the whole of human participation in reality as active subjects. Revelation, as a form of reason, is reason ecstatic. In this sense, ecstasy means "standing outside one's self." In this state of ecstasy, human reason points outside of itself to a state of mind that is extraordinary in the sense that the mind "transcends its ordinary situation and the subject-object structure."<sup>11</sup> Being in a state of revelation is the deepest form of participation in Being because we are actively involved in the mystery of Being.

Tillich lays out many possible mediums of revelation: nature, history, people, and the concept of the word (written, spoken, or otherwise—meant to symbolize the biblical "Word" and the Logos principle). All of these mediums are symbols, and are therefore imperfect bearers of revelation. In this sense, revelation is a form of *via negativa*; all revelation is both an affirmation and negation of what is being revealed. With all mediums of revelation, both everything and nothing, and everyone and no one are bearers of revelation. This is because nothing can perfectly manifest the mystery of Being, but everything is a possible bearer in a finite and imperfect sense. This aspect of Tillich's notion of revelation laid the foundation for his turn toward religious pluralism late in his life.

Before we deal with these aspects of Tillich's later thought, we should examine the other elements of his theology and their Kantian foundations. Since so much has been said of Tillich's theology,<sup>12</sup> I

---

10 Tillich, *Systematic Theology*, vol.1, 72.

11 Tillich, *Systematic Theology*, vol.1, 111.

12 See Charles W. Kegley and Robert W. Bretall, *The Theology of Paul Tillich* (New York: The Macmillan Company, 1952); Allen R. Killen, *The Ontological Theology of Paul Tillich* (J. H. Kok and N. V. Kampen, 1956); Adrian Thatcher, *The Ontology of Paul Tillich* (Oxford: Oxford University Press, 1978); Ian E.

shall here briefly sketch the relevant elements of his theological system as they relate to Kant's criteria for rational religious faith. First, Tillich sees humanity as being in a state of evil, or sin.<sup>13</sup> He expresses this mainly through the existential concept of estrangement. For Tillich, this is the essential problem of the human situation of finitude. He sees conversion<sup>14</sup> and banding together in the expectation of the coming of the Kingdom of God<sup>15</sup> as the solution to the problem. Tillich's notion of conversion is very similar to Kant's in many ways. While Kant sees conversion as taking on the good disposition of the prototype as we aim at moral perfection, Tillich, being a Christian theologian, sees conversion as based on the Christ event in the quest for the New Being of humanity. Even though Tillich's theology contains these distinctly Christian elements, he still makes the turn toward religious pluralism. This is possible because of his view of revelation. The main theme running throughout Tillich's work is that of ultimacy or universality. Tillich sees having faith or being religious as being in a state of ultimate concern. This, for Tillich, has a double meaning. Both the concern itself and the object of concern must be ultimate. For something to be truly ultimate, it must be universal. For Tillich, the ultimacy of God and the universality of reason lead to the universality of revelation. While Tillich sees the Christ as being universal as the final revelation, there are other forms of revelation that participate in the universal nature of reason.

We have seen that revelation is, for Tillich, a much broader concept than it is for most theologians; Tillich sees it as the ecstatic element of reason whereby we are able to grasp and be grasped by the meaning of our Being. We have already seen that Tillich's conception of revelation has many aspects. Because of this multifaceted nature, revelation can also have many forms. Tillich claims that the "divine Spirit or God, present to man's spirit, breaks into all history in revelatory experiences which have both a saving and transforming character ... [through] ... universal revelation and the idea of the holy."<sup>16</sup> This is possible because both hu-

---

Thompson, *Being and Meaning: Paul Tillich's Theory of Meaning, Truth and Logic* (Edinburgh: Edinburgh University Press, 1981).

13 See Tillich, *Systematic Theology*, vol.2.

14 See Tillich, *Systematic Theology*, vol.2.

15 See Tillich, *Systematic Theology*, vol.3.

16 Tillich, *Systematic Theology*, vol.3, 140. See also, Paul Tillich, "Appendix A: The Category of the 'Holy' in Rudolf Otto" and "Appendix B: Thinkers of Today: Rudolf Otto—Philosopher of Religion" in Firestone, *Kant and Theology at the Boundaries of Reason*.

manity's need for salvation and its participation in Being are universal. Of this, Tillich says, "[t]he universal quest for the New Being is a consequence of universal revelation ... [and] ... the different forms in which the quest for the New Being has been made are fulfilled in Jesus as the Christ."<sup>17</sup> We see from these passages that Tillich viewed the universal character of reason, revelation, and participation in Being as the foundations for the validity of many different types of religious experiences, provided they are transformative and salvific. For Tillich, religious experiences from many of the world's religions can have these characteristics because they are fulfilled in Jesus as the Christ. Here we see that Tillich is working with a type of fulfillment model of religion, where the religious experiences of all humanity, while authentic in themselves as bearers of meaning through participation in Being, gain their ultimate meaning from participation in God through fulfillment by the Christ event. Because of this, Tillich is able to be both an apologetic Christian theologian<sup>18</sup> and a religious pluralist. However, in 1963 after the completion of his three volume *Systematic Theology*, Tillich has moved to seeing religion as ultimate concern as the "criterion by which to judge the concrete religions."<sup>19</sup> My qualm is not with Tillich's system in general or his notion of ultimate concern in particular, but only with ultimate concern as the main criterion for judging religions. I contend that Tillich's theology of religions, in order to be authentically Kantian, needs to have the Kantian elements of rational religious faith as its main criteria.

#### 4. Kant's Rejoinder to Tillich: Rational Religious Faith as Transcendental Inclusivism

In this section, I will attempt to do something that may seem very strange: I will attempt to use Kant in order to offer a more sure foundation for theologians to bolster their theology of religions. I call this

---

17 Tillich, *Systematic Theology*, vol. 2, 89.

18 See Tillich, *Systematic Theology*, vol. 1.

19 D. Macenzie Brown, *Ultimate Concern: Tillich in Dialogue* (New York: Harper & Row, 1965), 4; See also, Paul Tillich, *Christianity and the Encounter of the World Religions* (New York: Columbia University Press, 1963).

transcendental inclusivism<sup>20</sup> because, following Palmquist, I see Kant as giving the necessary conditions any religion must meet in order to be included in the category of rational religion.<sup>21</sup> Kant sees these necessary conditions as the elements of rational religious faith. In laying out the elements of rational religious faith in *Religion* Kant goes some distance toward answering the question of the nature of humanity. Firestone has also presented these elements in several places,<sup>22</sup> so I will briefly sketch them rather than rehash others' arguments here. In Book One of *Religion* Kant claims that humanity is evil by nature; he goes on in Book Two to present the notion of grace through the prototype of perfect humanity;<sup>23</sup> finally, in Book Three, Kant turns to his vision of the Kingdom of God on Earth as he talks of the need for us to band together into "churches" in order to overcome radical evil. According to Kant, for any religion to be considered rational, it must contain, or at least have the resources to contain, these elements. Even though this initially makes Kant's philosophy seem as if it is nothing more than Christianity in disguise, a careful reading of *Religion* proves otherwise.<sup>24</sup> For Kant, the evil in human nature is a rational notion of radical evil instead of the doctrine of original sin. Also, the prototype is not a philosophical conception of Jesus; rather this is the rational conception of humanity as the goal we should aim at in our moral striving. Kant speaks of this not as an individual goal, but as a goal for the entirety of the human species; this is in keeping with the notions of the universality of radical evil in human nature and the need for all of humanity to undergo conversion

- 
- 20 Stephen Palmquist recommended the term "transcendental exclusivism" in relation to "empirical inclusivism."
- 21 This claim forms the backbone of Palmquist's exhaustive interpretation of Kant's *Religion* as a system of "religious perspectives" in Stephen R. Palmquist, *Kant's Critical Religion* (Aldershot, UK: Ashgate Publishing Ltd., 2000), esp. Part Three.
- 22 Chris L. Firestone, *Kant and Theology at the Boundaries of Reason*; "What Can Christian Theologians Learn from Kant?", *Philosophia Christi* 1 (2007); "Rational Religious Faith and Kant's Transcendental Boundaries", in Kevin Vanhoozer and Martin Warner (eds.), *Transcending Boundaries in Philosophy and Theology: Reason, Meaning and Experience* (Aldershot, UK: Ashgate Publishing Ltd., 2007).
- 23 See Nathan Jacobs, "Kant's Prototypical Theology: Transcendental Incarnation as a Rational Foundation for God-Talk" in *Kant and the New Philosophy of Religion*.
- 24 See Chris L. Firestone and Nathan Jacobs, *In Defense of Kant's Religion* (Bloomington: Indiana University Press, 2008) and Palmquist, *Kant's Critical Religion*.

to the good disposition of the prototype and band together to work toward the coming of the Kingdom of God on Earth.

While it is not completely clear from his writings that Tillich made the full turn to religious pluralism, most interpreters see him as doing so; and it is clear that he was at least heavily leaning in that direction. Moreover, we have seen the resources in his theology for making such a turn; also, with the notion of ultimate concern as the only religious criterion, he had nothing clearly holding him back. We have already seen that Tillich's view of Christianity meets Kant's criteria for rational religion; however, Tillich does not hold all religions up to this standard. Just because a religious experience has the effect of changing someone's life does not validate the tradition or religion in which the experience took place. Admittedly, Tillich never claims it does; however, his theology, as it stands, leaves us on a slippery slope at just this point. As I see Tillich's theology, there is no safeguard against adopting religious pluralism. The reason this is troubling is that, on this way of approaching religious diversity, assuming there is ultimate meaning and we are able to discover it, there is no way to determine what it is or how this is to be done.

For Kant, a religion must have the rational core necessary to deal with the problem of radical evil in order to answer the question of moral hope. Kant's arguments are aimed at finding a form of religious faith that can be justified rationally. This does not necessarily (1) exclude other types of faith from being, in many ways, legitimate, only from having the ability to claim rational justification, nor does it (2) deal with any of the individual adherents of any particular religion. In these ways, Kant's notion of rational religious faith is not an answer to the enduring question of religious pluralism, but only as it regards the rationality of a particular religion. As we have seen, Kant sees us as needing to postulate God, freedom, and immortality in order to have hope in our moral striving. Also, in light of the problem of radical evil he outlines in *Religion*, we need to have a conversion to the prototype of perfect humanity and then band together in ethical communities in the hope of the realization of the highest good. Without these core elements, Kant sees no way to meet the needs of reason. Some foundation is needed that is sufficiently able to meet these needs and answer these questions. It does not have to be Kant's necessarily. I say, along with Kant (6:10), that another solution is welcome, provided "a better one of the same kind can be found."

BOOK THREE:

EAST-WEST PERSPECTIVES  
ON CULTIVATING PERSONHOOD

PART IX

Mou Zongsan and Kant  
on Intellectual Intuition





## 43. Mou Zongsan's Critique of Kant's Theory of Self-Consciousness in the First *Critique*

Wen-berng Pong

### 1. Kant and Mou on Self-Consciousness and Knowledge of Objects

“Know Yourself” is an ancient Greek aphorism inscribed in the forecourt of the temple at Delphi. But how can we know ourselves? Can we apply the same method to know ourselves as to know objects? Is this not in vain, if the self does not exist at all? It seems to me that Kant puts two different kinds of knowing, knowing objects and knowing ourselves, on the same boat. He argues in the first *Critique* the following thesis: if it is possible for us to know any object in experience, we have to presuppose self-consciousness (transcendental apperception) as its possible condition. Kant begins his theory of self-consciousness with the famous statement: “It must be *possible* for the *I think* to accompany all my representations.”<sup>1</sup> But even if the possibility of knowing objects presupposes the existence of self-consciousness, Kant does not really bind knowing objects and knowing ourselves together, but only combines knowing objects with ourselves. We do not need to have, in this case, any knowledge of ourselves. What happens if we turn the knowing activity to ourselves instead of to objects? Here we confront a lot of puzzles. How can we know our self-consciousness as an observed object by means of the same entity that is already involved in the knowing activity? Does Kant apply the same method to know objects and to know ourselves? This seems to be the case at first sight. We have roughly this impression from the following passage (4:451):

[A]ll representations which come to us involuntarily enable us to cognize objects only as they affect us and we remain ignorant of what they may

---

1 *Immanuel Kant's Critique of Pure Reason*. tr. N. K. Smith (London: Macmillan, 1961), B131.

be in themselves .... Even as to himself, the human being cannot claim to cognize what he is in himself through the cognizance he has by inner sensation ... although beyond this constitution of his own subject, made up of nothing but appearances, he must necessarily assume something else lying at their basis, namely his ego as it may be constituted in itself.

But if we take the category of substance into consideration, it seems not to be the case that knowing of objects and knowing of ourselves are parallel. According to Kant, we can apply substance to constitute the object of experience, without extending our knowledge to thing-in-themselves; but we cannot apply the same category to constitute ourselves as having the same theoretical status as objects of experience, even if we do not extend our knowledge to “ourselves-in-themselves”. He writes in A349–50:

So far from being able to deduce these properties merely from the pure category of substance, we must, on the contrary, take our start from the permanence of an object given in experience as permanent .... The “I” is indeed in all thoughts, but there is not in this representation the least trace of intuition, distinguishing the “I” from other objects of intuition.

Kant introduces the principle of self-consciousness in order to solve the problem of knowledge of objects, but the result is different for the problem of knowing oneself. We have the following famous sentence (B157): “[In] the transcendental synthesis of the manifold of representations in general, and therefore in the synthetic original unity of apperception, I am conscious of myself, not as I appear to myself, nor as I am in myself, but only *that* I am.” Kant seems to maintain in the above statement that the “I” has three different forms: appearing I, transcendental I, and I in myself. In order to interpret this passage in the right direction one may take an example from Kant’s theory of objects. As Allison suggests, the distinction between appearance and thing-in-itself can be interpreted in two ways, as either a two-world theory or a two-aspect theory.<sup>2</sup> According to the former, appearance and thing-in-itself are ontologically different entities; but according to the latter, they are two aspects of one and the same object with respect to different knowing capabilities. Allison prefers the latter interpretation rather than the former. One could try to interpret Kant’s theory of self-consciousness

---

2 H. E. Allison, *Kant’s Transcendental Idealism: An Interpretation and Defense* (New Haven and London: Yale University Press, 2004), 3–7. Cf. also K. Ameriks, “Recent Work on Kant’s Theoretical Philosophy”, *American Philosophical Quarterly* 19 (1982), 1–24.

in an analogical way to his theory of objects and ask: does Kant have in mind a three-I theory or a theory of three-aspects of one and the same I?

As Allison did in the two-aspect theory of objects, Mou Zongsan holds (1) that Kant should have a “three-aspect theory” of self-consciousness in mind; but he argues (2) that a “three-I theory” would be justifiable. I shall begin this essay with the question whether Mou’s thesis has enough textual support, and then examine how Mou criticizes this thesis in order to justify his second thesis. One should remember that even if Mou falsely interprets Kant in the first thesis, his critique still proposes some new perspectives in the theory of self-consciousness: instead of total ignorance of the “I in myself”, we can have some self-knowledge of the “I in myself” in a “three-I” theory. Finally I will argue on the side of the “three-aspect” theory and see if Mou’s critique is defensible.

## 2. Kant’s Three-aspect Theory of the Same I

Since Mou’s critique of Kant’s theory of self-consciousness is founded on a three-aspect theory of the I, one should ask how far this premise can be confirmed. According to a two-aspect theory of objects, appearance and thing-in-itself belong respectively to different uses of our knowing capabilities: appearance by the combination of sensibility and thing-it-itself by the understanding. Do we use the same capability to cognize ourselves as appearance as we use to cognize objects as appearance, and likewise with ourselves and objects as they are in themselves?

What is an “appearing I”? Since we ascribe some properties to an appearing object, insofar as it affects us as knowing subject through outer sense—for example, we ascribe the red and the scent to the rose, insofar as the red and the scent are given to us (B70)—we have to ascribe some properties in the same way to the “appearing I”, insofar as the latter affects us; for example, Obama ascribes pleasure in his presidential campaign to himself as “appearing I”, insofar as the latter affects the former. For Kant we know the appearing I exactly as we know the appearing object (B156):

[Since] we admit that we know objects only in so far as we are externally affected, we must also recognize, as regards inner sense, that by means of it we intuit ourselves only as we are inwardly affected by ourselves, in other

words, that, so far as inner intuition is concerned, we know our own subject only as appearance, not as it is in itself.

In the case of objects as appearance I can well imagine that they affect the subject through outer sense in some way, but in the case of intuition of ourselves, how can the I as appearance affect the subject? It is quite strange to say that the knowing subject has knowledge of itself as appearance insofar as the appearing self affects itself as knowing subject; furthermore, beyond this self-relation of knowing subject and appearing I, there still exists I in myself, if one abstracts from this self-relation. If Kant thinks the same relation exists between knowing an object and knowing oneself, and maintains a two-aspect theory instead of two-world theory, then the distinction between the appearing I and the I in itself lies in a different use of the knowing capability: the appearing I is the I from the aspect of sensibility (i. e., inner sense), and the I in itself is the I from the aspect of understanding. Since the knowing subject does not have any positive content, knowable through understanding alone, the I in itself is an empty concept for the knowing subject, just as the case in the two-aspect theory of objects. But if we take Kant's explanation of transcendental apperception into consideration, the aspect of I in itself is difficult to distinguish from that of apperception, because they both belong to the understanding (B155, my emphasis):

How the "I" that thinks can be distinct from the "I" that intuits itself (for I can represent still other modes of intuition as at least possible), and yet, as *being the same subject can be identical with the latter*; and how, therefore, I can say: "I, as intelligence and thinking subject, know myself as an object that is thought, in so far as I am given to myself [as something other or] beyond that [I] which is [given to myself] in intuition, and yet know myself, like other phenomena, only as I appear to myself, not as I am to the *understanding*" ....

It seems the two aspects of sensibility and understanding in respect of "I" belong to the relation between the appearing I and apperception, rather than to the relation between the appearing I and I in myself. A problem arises if we ask: does apperception lie inside or outside of the understanding? In the two-aspect theory of appearance and thing-in-itself the latter seems to lie outside the understanding, but in case of the appearing I and apperception the latter seems to lie inside the understanding: "The unity of apperception in relation to the synthesis of imagination is the understanding; and the same unity, with reference to the transcendental synthesis of the imagination, the pure understand-

ing" (A119). That is to say, as a part of understanding, apperception can have on the one hand different kinds of synthesis, and can be itself on the other hand the object of understanding. The apperception can function as a subject to synthesize other representations and at the same time as an object in the understanding to be synthesized. Both belong to the activity of understanding, but it seems the first is transcendental and second in the time condition. Is the appearing I an aspect known purely through self-intuition of the I, without the use of the "I" as a concept in the understanding, or is it an aspect of the concept "I" in combination with self-intuition? We know objects only in appearance, insofar as we apply concepts with the corresponding intuitions; for example, I apply the concept of rose in the understanding to the intuition of several properties (e. g., red and scent) in the sensibility in order to cognize the object in appearance: a real rose. Likewise, I should apply the concept of "I" in the understanding to the intuition of several properties (e. g., pleasure) in the sensibility in order to cognize an "appearing I": Obama. We have here the ambiguity of the "appearing I". It seems to me that sometimes "appearing I" signifies an "I" purely by means of self-intuition (see B156, above; cf. also B68), and sometimes signifies a concept of "I" plus some properties of inner intuition (B158):

Just as for knowledge of an object distinct from me I require, besides the thought of an object in general (in the category), an intuition by which I determine that general concept, so for the knowledge of myself I require, besides the consciousness, that is, besides the thought of myself, an intuition of the manifold in me, by which I determine this thought.

In spite of this ambiguity of the "appearing I", we can at least be certain that the appearing I and apperception are two different aspects of the same subject, but it still lacks a third aspect of the same subject: the aspect of the I in itself. We have no evidence in the text to support this thesis, but we do find some hints of this possibility (B158–9, my emphasis):

I exist as an intelligence which is conscious solely of its power of combination; but in respect of the manifold which it has to combine I am subjected to a limiting condition (entitled inner sense), namely, that this combination can be made intuitable only according to relations of time, which lie entirely outside the concepts of understanding, strictly regarded. Such an intelligence, therefore, can know itself only as it appears to itself in respect of an intuition which is not intellectual and can not be given by the understanding itself, not as *it would know itself if its intuition were intellectual*.

We can summarize from the above description that transcendental apperception is a faculty in the understanding that produces the concept of “I”. In contrast with apperception, the appearing I and the I in itself correspond, apart from the concept of “I”, to types of intuition that lie beyond the understanding, the former with sensible and the latter with intellectual intuition. If one can confirm that in Kant’s mind the “I” seen through intellectual intuition is an aspect of the same subject that can also be seen as the appearing I through sensible intuition, then the premise of Mou’s critique under the three-aspect theory is confirmed. Kant does not say this explicitly, but we can find some indirect textual support (A346/B404, my emphasis):

We can assign no other basis for this teaching than the simple, and *in itself completely empty, representation “I”*; and we cannot even say that this is a concept, but only that it is a bare consciousness which accompanies all concepts. Through this I or he or it (the thing) which thinks, nothing further is represented than a transcendental subject of the thoughts = X. It is known only through the thoughts which are its predicates, and of it, apart from them, we cannot have any concept whatsoever, but can only *revolve in a perpetual circle*, since any judgment upon it has always already made use of its representation.

Kant argues here that epistemologically the existence of transcendental apperception or bare consciousness can only be indirectly concluded from the thoughts as its predicates. It is in itself completely empty. Any judgment of I in itself would revolve in an unavoidable circle. If Kant had a theory of three different Is instead of three aspects of the same I, such a circle would not occur, since the judgment of I in itself does not need to make use of transcendental apperception. In other words, Kant should have in mind that we could have *a priori* self-consciousness, but without *a priori* knowledge of the self (i. e., of the I in itself): “Accordingly I have no knowledge of myself as I am but merely as I appear to myself. *The consciousness of self is thus very far from being a knowledge of the self*” (B158, my emphasis). Consciousness of an “I in itself” is not the same as knowing what it is. Put simply, Kant maintains that we have self-consciousness without self-knowledge in a transcendental sense. Kant’s whole argument seems to run as follows. Should the appearing I, apperception and the I in itself be three different aspects of the same I, then one should have transcendently only self-consciousness without self-knowledge, There are two reasons: firstly, we do not have any corresponding *a priori* intuition in order to have cognition of the “I in itself”; secondly, even if we have such intuition, no

matter what the content of such self-knowledge is, it would incur an avoidable circle. But it seems we here face another circle: If I could not have any knowledge of the "I in itself", how could I know that the appearing I, apperception and the I in itself are three different aspects of the same I? This is the watershed between Kant and Mou. It is very strange to notice that in order to keep the premise of the three-aspect theory and avoid the possible circles, Kant is forced to maintain the thesis of self-consciousness without self-knowledge; Mou argues, on the contrary, that Kant's premise of the three aspect theory cannot avoid some other contradictions, so that we had better adopt the three-I theory instead of the three-aspect theory, for in this way we can extend self-consciousness to self-knowledge.

### 3. Mou's Three-I Theory

Mou begins with the premise that Kant should have a three-aspect theory of the I. Kant seems to argue in the Paralogisms that this theory should be the most logically consistent (i. e., defensible) theory. In his *Intellectual Intuition and Chinese Philosophy* Mou argues (1) that Kant's three-aspect theory incurs various contradictions if we apply it to Kant's theory of self-consciousness as a whole, and (2) that the alternative, the theory of the three Is, can avoid those contradictions.

Mou summarizes his interpretation of Kant's theory of self-consciousness as follows:

According to Kant, "I think" and "I am" are identical, one does not infer from "I think" to "I am". But the "I am", which is identical with "I think", is merely a simple representation in thinking, the so-called "am" does not mean the "being" in the sense of appearance, nor in the sense of thing-in-itself, that is, its being is undetermined. Accordingly, I have three aspects, that which is indicated by (is conscious of) the apperception is the simple "I am"; that by sensible intuition is the I as appearance; that by intellectual intuition is the I as in itself. Actually they are all of the same I.<sup>3</sup>

Three aspects means one can use three different cognitive capacities separately to approach the same I as their object. The appearing I is the first aspect, under sensibility (inner sense). We can imagine that we inwardly

---

3 Zongsan Mou, *Intellectual Intuition and Chinese Philosophy* (智的直覺與中國哲學, *Chih te chih chueh yü chung-kuo chih süeh*) (Taipei: The Commercial Press, 1971), 163; my translation.



intuit ourselves at a certain point of time without the help of any concept or thought, so that every mental state would connect with this intuited self. But since the intuited self is an object to be cognized in the inner sense, we have to ask: who in this situation is the cognizer (cognitive subject)? Is it the same as the intuited self? How could an intuited object turn out to be a cognitive subject? In B69 Kant agrees that this is a difficulty common to every theory of self-consciousness, but in a later footnote (B156) Kant argues on the contrary that understanding it is no difficulty.

In the second aspect we have a concept of I as the object to be cognized in the understanding by the same cognitive subject, or maybe by some other subject. According to Kant, this concept of I as the object of understanding is “a thought, not an intuition” (B157). Kant argues further that under such circumstances we can have self-consciousness, but without any self-knowledge *a priori*, since in order to have such self-knowledge some non-sensible intuition would be required that cannot be found in human cognition.

In connection with such non-sensible intuition consists the third aspect of the I, whose existence on the one hand is confirmed by Kant, yet on the other hand lies for him beyond any human cognition. I exist as an intelligence that is conscious solely of its power of combination; but in respect of the manifold that it has to combine, I am subjected to a limiting condition (entitled inner sense): that this combination can be made intuitable only according to relations of time, and this lies entirely outside the concepts of understanding, strictly regarded. Such an intelligence, therefore, can know itself only as it appears to itself in respect of an intuition that is not intellectual and cannot be given by the understanding itself, not as it would know itself if its intuition were intellectual (B158–9).

The self as intelligence can only be understood as a power of combination. For Kant there are four different kinds of object: in respect of I, either as appearance or in myself; in respect of a thing either as appearance or in itself. As an empty power of combination, intelligence requires something to combine in order to have cognition of the combined object; if human beings could have intellectual intuition, then they would have cognition of I in myself as well as the thing in itself. But one may ask: does the I in myself that intelligence is unable to combine lie inside or outside intelligence? “About this, it seems that Kant is not aware of it as a problem, [for he] eventually does not express it clear-

ly.”<sup>4</sup> In spite of Kant's unawareness Mou maintains that Kant should hold that I in myself lies inside intelligence. This implies that Kant's theory of self-consciousness should be a three-aspect theory instead of a three-I theory. If Kant really considers three-aspect theory as true, so argues Mou, then he can not keep the following thesis in the Paralogisms without contradiction: “‘Every thinking being is, as such, a simple substance’ is a synthetic a priori proposition” (B410). Mou argues that whether this is a synthetic *a priori* proposition or not depends on which field the concept of thinking being applies to. If “every thinking being is, as such, a simple substance” is a synthetic *a priori* proposition, then it is only valid as applied to the objects of experience, so that the thinking being as an “I” would become an object of sensible intuition, and “thinking being” would be determined by the categories. So understood, this kind of metaphysical determination would be a “determination of immanent metaphysics”. But insofar as we use the expression “synthetic a priori proposition” to refer to this thinking being as the transcendent (true) I, it is impossible for us to mean it is a sensible object. If we can presuppose an intellectual intuition, as a means to intuit this transcendent I, then the presupposed intellectual intuition requires neither categories, nor synthesis, so that “every thinking being is, as such, a simple substance” would not be a synthetic *a priori* proposition. This makes Kant's assertion self-contradictory.<sup>5</sup>

Since this contradiction is due to the premise of the three-aspect theory, Mou suggests adopting the alternative, three-I theory in order to avoid this contradiction. I conclude by quoting the argument Mou presents in defense of this alternative:

If we want to remove this contradiction, that is, to separate the synthetic a priori proposition from the I in myself in order not to let them tangle up, then the only way is to divide this concept of “I” into three different Is, rather than one I alone by means of “thinking subject”. Because Kant has in mind only one I, and because different aspects come from different approach to it, there are a lot of tangles. If we divide it into three Is, then the thinking subject is only a formal I, the logical I, the cognizing subject; by means of it, we spring into a transcendent true I as its basis and substratum. There is a certain distance between this transcendent true I and the logical I (cognizing subject). For the latter, we cannot constitute any “metaphysical determination”; that is, basically we do not have a synthetic a priori proposition at all; nor for the former can we say it is a synthetic a priori

---

4 Mou, *Intellectual Intuition*, 262.

5 Mou, *Intellectual Intuition*, 173.

proposition, because the true I is a simple noumena, an self-identical indestructible noumena, and so on; they are but analytic terms. It would manifest itself only by means of the lightening of intellectual intuition. We can regard it as a synthetic a priori proposition only insofar as we sense this true I by means of sensible intuition and determine it by means of the categories. The I that is determined to be such a synthetic a priori proposition is only an appearing pseudo-I. It is neither cognizing subject (logical I), nor the transcendent true I. The confusion comes because we take the cognizing subject (thinking subject) to be the same as the transcendent true I and so we incur those tangles and contradictions regarding “synthetic a priori proposition”: because it sways between two kinds of metaphysical determination, there arises such tangles; because transcendent true I cannot be considered to be synthetic a priori, there arises such contradictions. If we separate all three Is, those difficulties would disappear.<sup>6</sup>

The categorial determinations of simple and substance are applied *a priori* to the concept of thinking being. For Mou the thesis is only valid in the categorial determination of I as appearance, but not valid for both the I in myself and apperception. For me, this thesis cannot be applied to all three of these aspects of I. For apperception, this application cannot result in any metaphysical determination: “The analysis, then, of the consciousness of myself in thought in general, yields nothing whatsoever towards the knowledge of myself as object. The logical exposition of thought in general has been mistaken for a metaphysical determination of the object” (B409). For the appearing I, we cannot even apply the category of substance in order to constitute the I as an object of experience, as with other objects (A349–50). For the I in myself, we can indeed obtain some metaphysical determination if our intuition were intellectual. However, this would not be the case if we do not have such intuition.

---

<sup>6</sup> Mou, *Intellectual Intuition*, 173–4.

## 44. Mou Zongsan and Kant on Intellectual Intuition: A Reconciliation

*Mihaela C. Fistioc*

### 1. Introduction

Mou Zongsan famously criticizes Kant<sup>1</sup> for denying human beings what Kant calls intellectual intuition [*intellektuelle Anschauung*] (5:409 [§77]). In this essay I try to show that, in spite of their apparent opposition, Mou Zongsan's and Kant's positions on this issue can be reconciled. Kant discusses the notion of intellectual intuition in the *Critique of Pure Reason*. He takes up this notion again in the *Critique of the Power of Judgment*, this time more often under the name of "intuitive understanding" [*intuitiver Verstand*] (5:406 [§77]).<sup>2</sup> The paradigmatic example of intuitive understanding is God; by contrast, human beings have what Kant calls a "discursive" understanding.<sup>3</sup> An intuitive understanding grasps everything directly and in full. A discursive understanding, by contrast, grasps everything indirectly, through concepts, and only to a certain degree. An intuitive understanding is infallible, while a discursive understanding can make mistakes.

Importantly, in the third *Critique* Kant characterizes the two types of understanding in terms of purposes. An intuitive understanding sees the purpose, the rationale of everything to the core. A discursive understanding, while lacking full grasp of every purpose in nature, nevertheless proceeds *as if* everything had a purpose. So, a discursive understanding *imitates* an intuitive one. This imitation is the reconciliation I see between Mou Zongsan's and Kant's positions. While Kant does indeed claim that human beings have only a discursive understanding, he also

---

1 Mou Zongsan, *Phenomena and Noumena* (Taipei: Student Book Company, 1975), 3.

2 See §77 for the use of these two terms interchangeably.

3 Parts of the account to follow have appeared previously in Mihaela C. Fistioc, *The Beautiful Shape of the Good: Platonic and Pythagorean Themes in Kant's Critique of the Power of Judgment* (New York and London: Routledge, 2002).

characterizes such a discursive understanding as attempting to approximate an intuitive one. Human beings, one might say, have “as if” intuitive understanding.

## 2. Discursive versus Intuitive Understanding

Here is how Kant defines the notion of end, or purpose [*Zweck*] in his third *Critique* (5:220 [§10]): “an end,” he says, “is the object of a concept insofar as the latter is regarded as the cause of the former (the real ground of its possibility); and the causality of a **concept** with regard to its **object** is purposiveness (*forma finalis*).” So that is Kant’s definition. But what are we to make of it? Let me try to explain this by means of an example.

In my study, back in New Haven, Connecticut, I have this funny object on my desk. It is a pile of plastic pieces: a dark gray square plastic slab at the bottom, another light gray square plastic slab on top, a coiled plastic string, and another longish plastic slab connected with the coil and lying on top of the light gray square. It is an object I can touch and move around. I bought it in Cambridge, Massachusetts. And now let me offer a different description: it is the result (or the object, as Kant would put it) of an idea in the mind of an engineer at Bell Phone; the engineer had in mind a tool that would allow people like me to communicate with other people all over the world. By now you know that what I have before me is a telephone. Suddenly, the longish plastic slab on top is not just a slab, but a receiver; it can capture my voice and transmit it down the coiled line, through the circuit inside the phone, then on to the line in the wall. To be able to treat it as a telephone, I must view this object as the effect of the engineer’s idea; this idea is what Kant, with Aristotle, would call the “final cause” of the object. As soon as I take that idea away, I am left with an odd arrangement of plastic pieces, produced on some assembly line.

In the *Critique of the Power of Judgment*, Kant discusses at length this notion of end in the context of the distinction between a mechanistic and a teleological approach to the exploration of nature.<sup>4</sup> The mechanistic approach, Kant explains, proceeds from the parts to the whole, and the whole is only a sum of parts. (It is this mechanistic approach that I had in mind when I first described the object on my desk as a mere ag-

---

4 See the Dialectic of the Teleological Judgment, especially §§76–8.

glomeration of plastic pieces.) The design, or teleological approach, the approach in terms of ends, proceeds from the whole to the parts; in this case the whole is not the sum of parts, but rather what first makes sense of how the parts function and interact, what explains their *raison d'être*. (In my example, this was the telephone with all its functionality, as designed by the engineer.) But to explain why something is where it is, and the way it is, means to make it intelligible; this is why the notion of design is commonly connected with the thought of intelligence as its source.

Kant is concerned with the application of the distinction between a mechanistic and a teleological approach to our experience of nature, and his question is this: in our approach to nature, do we human beings, as discursive understandings, first know the parts and compound them into an aggregate or do we first know the design of the whole? His answer is that a discursive understanding is forced by its own nature to start from the parts, rather than from the whole of nature as intelligent design. For design has to do with an intelligence as author and, while it is true that we ourselves are intelligences, it is just not the case that we create the content of our experience. An intuitive understanding, by contrast, creates its objects of experience according to its own ideas, which means that such an understanding does start from design.<sup>5</sup> So a discursive understanding naturally thinks, as it were, in terms of plastic slabs, whereas an intuitive understanding thinks in terms of telephones, or rather only one giant telephone, the world, which it creates as it thinks it. (In this context “thinking” has to do with the origin and kind of knowledge

---

5 One may find it puzzling to talk of God as thinking in terms of ends; only we human beings, one might point out, set up goals and then proceed to work towards them. But this need not be the way we interpret the connection between God and ends. God does not have an intermediate step between thought and action, nor does It have to approximate towards the implementation of a design. Rather, whatever It thinks comes into being instantaneously or rather simply *is*; and whatever It thinks shows intelligence through and through or, put differently, is fully designed. As designed, God’s work, the world, looks to us as having an end; we cannot help but make the connection between design and a creative intelligence. Kant himself, it should be said, speaks of God in terms of ends: “we will conceive of this original being ... as *omnipotent*, so that he can make the whole of nature suitable for this highest *end*” (5:444 [§86]; second emphasis mine). Kant believes we cannot help but think of God in human terms, particularly in moral terms. Indeed, he believes that it is precisely this moral dimension that can give real content to the concept of God.

each type of understanding can possess: a discursive understanding can gain knowledge only by starting from the slabs, bottom-up, while an intuitive understanding starts from the overall picture and views the world top-down.) Here is Kant's way of describing the distinction between the two types of understanding (5:407 [§77]; italics mine):

Now, however, we can also conceive of an understanding which, since it is not discursive like ours but is *intuitive*, goes from the **synthetically universal** (of the intuition of a whole as such) to the particular, i. e., *from the whole to the parts*, in which, therefore, and in whose representation of the whole, there is no **contingency** in the combination of the parts, in order to make possible a determinate form of the whole, which is needed by our understanding, which *must progress from the parts*, as universally conceived grounds, to the different possible forms, as consequences, that can be subsumed under it. In accordance with the *constitution of our understanding*, by contrast, a real whole of nature is to be regarded only as the effect of the concurrent moving forces of the *parts*.

### 3. God's Mind as Standard

This is, however, not the whole story, and here I come to Kant's possible answer to Mou Zongsan's criticism. While we ourselves do not have an intuitive understanding, we can nevertheless mimic it. We may well be forced to start from the parts, but we do not have to regard the parts as a mere agglomeration. We can also try to figure out the purpose of the parts viewed as a whole. We can try to outguess a putative God—in other words, we can take God's mind as our standard. Suppose, we tell ourselves, that this particular natural object were not accidental, the way it looks to us. How could it fit into the whole of nature such that it would be seen as necessary rather than haphazard? At every level of experience, this line of thought leads us to see the particular as one part of a larger whole, as serving a function within a complex. But to see something as possible through an idea coming from intelligent design, as the effect of such intelligent design, is, as shown in the telephone example above, to think in terms of ends. So, while a discursive understanding naturally starts by thinking in mechanistic terms, it also has the capacity, and therefore the option, to think in terms of ends (5:407–8 [§77]; English italics mine):

Thus if we would not represent the possibility of the whole as depending upon the parts, as is appropriate for our discursive understanding, but would rather, after the *model* [*Maßgabe*] of the intuitive (archetypal [*urbil-*

dlichen]) understanding, represent the possibility of the parts (as far as both their constitution and their combination is concerned) as depending upon the whole, then, given the very same special characteristic of our understanding, this cannot come about by the whole being the ground of the possibility of the connection of the parts (which would be a contradiction in the discursive kind of cognition), but only by the **representation** of a whole containing the ground of the possibility of its form and of the connection of parts that belong to that. But now since the whole would in that case be an effect (**product**) the **representation** of which would be regarded as the **cause** of its possibility, but the product of a cause whose determining ground is merely the representation of its effect is called an *end*, it follows that it is merely a consequence of the particular constitution of our understanding that we represent products of nature as possible only in accordance with another kind of causality than that of the natural laws of matter, namely in accordance with that of *ends* and *final causes* .... And further, it is not at all necessary here to prove that such an *intellectus archetypus* is possible, but only that in the contrast of it with our discursive, image-dependent understanding (*intellectus ectypus*) and the contingency of such a constitution we are led to that idea (of an *intellectus archetypus*), and that this does not contain any contradiction.

Note the last sentence. As beings with discursive understanding, we do not create the world as such; rather, we only picture the world in images, and so we re-present it. The word “archetypus” comes from the Greek adjective ἀρχέτυπος, meaning “first-molded as a pattern or model” (τὸ ἀρχέτυπον is, among other things, the figure on a seal), while “ectypus” comes from the adjective ἐκτύπος, meaning “formed in outline or worked in relief”.<sup>6</sup> These etymologies bring out the distinction between an intuitive and a discursive understanding: an intuitive understanding gives the original pattern and so creates it in an immediate way, while a discursive understanding takes on the shape of the original or paints pictures of the original and so, to the extent that it can, re-creates the original pattern. The important point that Kant himself stresses in the last sentence of the passage is that an *intellectus ectypus* has a need for images. We human beings, it turns out, make sense of the world in terms of images taken as standards.

So far I have spoken, in a rather general manner, of thought in terms of ends. It is now time to introduce a further distinction, which lies at the heart of Kant’s project in the *Critique of the Power of Judgment*: the distinction between an end or a purpose [*Zweck*] and purposiveness

---

6 *A Greek-English Lexicon*, compiled by Henry George Liddell and Robert Scott, revised by Henry Stuart Jones, ninth edition (Oxford: Clarendon Press, 1940 [new supplement 1996]).



[*Zweckmäßigkeit*]. As we saw earlier, in Kant's view an end is the instantiation of an idea that originated in the mind of an intelligent being, such as God or ourselves. In this light, we can easily regard artifacts, such as telephones, as ends, since they are the result of ideas in the mind of human beings. However, we cannot do this with nature, either in part or as a whole, since we do not have *knowledge* of a mind behind it, such as God's. All we can say of objects or complexes of objects in nature is that they *look* designed, that they appear to accord with, to be in the style of [*mäßig*] an end or purpose<sup>7</sup> [*Zweck*] and so they show purposiveness [*Zweckmäßigkeit*].

Part of what motivates Kant's discussion in the Dialectic of the *Critique of the Power of Judgment*, where the passage quoted above appears,<sup>8</sup> is his wish to show that we have no *knowledge* of a God behind nature, a God who sets the initial end (in the sense of purpose) of the world. Rather, the idea of such a God is only a regulative principle of our thought about nature. It is a model [*archetypus*] or a standard [*Maßgabe*] that we ourselves come up with in our endeavor to make sense of the world. But, while we cannot prove that various aspects of nature have precise ends set by such a God, it is the case that purposiveness, as the property of looking designed, is an all-encompassing feature of our experience of nature.

There is a sense in which the figure of God towers over the third *Critique*. On Kant's picture, while we are not God, we can, and indeed do, imitate God's way of thinking: we work painstakingly at seeing things from the top, in their interconnections, the way God sees things effortlessly. While we do not create the world, we can re-trace the steps of an imagined Creator and thereby, in effect, re-create the world on our own. We are not indifferent to the kind of life we lead: we want to cut the best figure in our short stay in this world. This leads to per-

---

7 I offer here "purpose" as an alternative translation for *Zweck* in order to bring out the connection with purposiveness [*Zweckmäßigkeit*]. I nevertheless agree with the point made by the editor of the text in support of the choice of "end" over "purpose": "purpose" for *Zweck* obscures the connection between Kant's aesthetics and his ethics" ("Editor's Introduction", in Paul Guyer [ed.], *Critique of the Power of Judgment*, trs. Paul Guyer and Eric Matthews [Cambridge: Cambridge University Press, 2000], xlviii).

8 Kant divides his *Critiques* into an Analytic and a Dialectic; as a rule, he describes in the Analytic various legitimate principles of the topic at hand; in the Dialectic he presents ways in which the principles just presented in the Analytic might be misused, such that they end up overstepping their legitimate domain.

petual experimentation and reflection (one reason Kant speaks of our *reflecting* power of judgment), even when we know that in fact it is impossible to have either the final truth or perfect happiness; as Kant puts it in the first *Critique*, we cannot help searching for the unconditioned (see e. g., A332/B389). To speak of God or the gods is ultimately to speak of a perfection. It is a perfection that beckons to us and invites us to make it our standard: the more we think like the gods, the closer to their state we ourselves are.

One might wonder whether Mou Zongsan would have been satisfied by this “as if” intellectual intuition. But here it is important to note that this “as if” is not a mere construct that can be eliminated at will. It is rooted, on Kant’s own account, in our very human nature: we respond with pleasure to the discovery of design—that is, of purpose. So our very nature pushes us to pick the rational, the intelligible way of looking at things. This is, indeed, the main message of the *Critique of the Power of Judgment*. Now this idea of our human nature guiding us, through feeling, down the right path has parallels in the Chinese philosophical tradition, and in particular in Confucianism. This is why I believe Mou Zongsan’s and Kant’s views can in fact be reconciled.<sup>9</sup>

---

9 I am grateful to Stephen Angle, Chun Ho Lai, Sanford Shieh, and Hanqing Yu for help in writing this essay. I am also grateful to Chung-ying Cheng for extensive comments during the conference panel on Mou Zongsan’s interpretation of Kant, where this essay was presented.

## 45. On Kant's Duality of Human Beings

Sammy Xia-ling Xie<sup>1</sup>

### 1. Thing in Itself

When referring to the duality (*zwiefache Art*) of human beings, Kant says that the human being is not only a phenomenon in the world of sensation (*Phänomen in der Sinnenwelt*), but also an intelligence in the world of understanding (*Intelligenz in der Verstandeswelt*).<sup>2</sup> The intelligence here should be regarded as thing in itself (*Ding* oder *Wesen an sich selbst*). Dealing with the thing in itself first, we will discuss intelligence later. Kant thinks the thing in itself is what affects (*affizieren*) us and gives us sensation, something behind phenomenon (A19/B33; 4:451). The thing in itself is included in the world of intelligence (*intelligibele Welt*) sometimes called by Kant “archetypal” nature. The world of sensation, also called “ectypal” nature, is the former’s counterpart (5:43).

Now we ask: what is the significance of Kant’s supposing supersensational nature, i. e., archetypal nature, as the source of sensation? Kant understands clearly that noumenon or thing in itself is merely a thought. He starts with the general inclination of the thinking of human beings, and says understanding will not be satisfied by sensible appearance and can think from phenomenon to noumenon. Kant adds that even the concept of appearance itself will lead us to the concept of noumenon, because the concept of appearance means a relation to something that is an object independent of sensation (A251–2). Now we have to refer to the transcendental object. The transcendental object

---

1 This essay was translated from Chinese with the assistance of Qiaozhen Wang.

2 Kant writes (4:457): “he must conceive and think of himself in this twofold way, which rests first on the consciousness of himself as an object affected by the senses, and second on the consciousness of himself as an intelligence, i. e., as independent of sensible impressions in the employment of his reason (in other words as belonging to the world of understanding).” We can translate the word “Intelligenz” into Chinese in different words, namely *lizhi* (理智), *zhili* (智力), *lingwu* (靈物), etc. Here, I use the word “*ti*”(體) to translate it because I think it should be regarded as substance.

is what equals X, a concept that has objective reality when sensible representation is connected to it (A109/B139). Sometimes Kant writes it as a transcendental "Gegenstand", but to put it definitely, it is only a thought about something in general, not a noumenon at all (A253: "Das Objekt, worauf ich die Erscheinung überhaupt beziehe, ist der transzendente Gegenstand, d. i. der gänzlich unbestimmte Gedanke von Etwas überhaupt. Dieser kann nicht das *Noumenon* heißen"; see also A109).

The transcendental object represents the unity of the manifold and is the *correlatum* of the unity of pure apperception. However, the concept of noumenon, or thing in itself (A254/B310, A256/B312), is a boundary concept, limiting the use of categories within the scope of sensible intuition. Kant adds a doctrine of positive noumenon on the premise that there is an intellectual intuition (*intellektuelle Anschauungsart*) (B307). However, to divide the world into one of sensation and the other of understanding in the sphere of nature must never be allowed (A255/B311).

To regard the thing in itself as a thought perhaps indicates that Kant has developed a kind of phenomenology. It is also possible that Kant only adheres to principles of transcendental logic and is not yet aware of phenomenology. It is logically coherent that we can think things as individual or things are plural. If we want to regard these as thing in itself, we have to use two categories, at least: the category of substance and the category of unity (or singular). However, in this case, we are no longer thinking of the thing in itself. If we do not use it in the plural sense, the thing in itself will be understood as a whole that will not be divided into things in particular. We can infer this logically. The requirement of the perspective of phenomenology is even more radical: canceling the possibility of individual things. As a result, we cannot say that there are objects that affect us; we can only say that we are affected.

There is a great thought in *GMM*. Kant says "the commonest understanding can form the concept of thing in itself by an obscure discernment of judgment which it calls feeling" (4:450–1). I regard this as a great thought because Kant has said how human beings know the thing in itself exists: the point here is that he attributes it to the power of feeling of the mind. It is a pity Kant does not develop this thought in his third *Critique*.

However, Kant regards the thing in itself and the world of understanding as the same immediately. He says of things in themselves

that, although we must admit they can never be known to us except as they affect us, we can come no nearer to them, nor can we ever know what they are in themselves. This must furnish a distinction, however crude, between a world of sense and the world of understanding.

To be sure, the living experiences of human beings make them recognize that the individual thing in itself really exists. People can regard their bodies as only phenomenon, only a concept; people can regard the relationship between father and son, between husband and wife, between employer and employee as well as between creditor and debtor as both noumenal and conceptual. But they cannot regard what they eat, or their spouse, or the sons and daughters they hug, as merely a concept (whether negative or positive) of phenomenon or intelligence or noumenon. Are these cases examples of what Kant means by “we can come no nearer to them”? Here it is necessary to adopt a concept of the thing in itself that is logically incoherent and to recognize that sensation is the result of the thing in itself affecting us. (If we want to avoid using the term, we have to adopt the perspective of phenomenology.) Whenever we begin to talk about philosophy, we have to use the concept of the thing in itself to refer to objects that cause sensation in us.

## 2. Intelligence as Thinking Subject

Intelligence refers to the subject in general, and primarily the thinking subject. Kant uses “intelligence” to refer to consciousness of the activities of spontaneity of the “I think”, self-consciousness, and mind (*Gemüt*; pure apperception).<sup>3</sup> In *CPR*’s B edition, §24, where Kant answers in particular how we understand “I”, the following sentence is typical: “*Ich*, als Intelligenz und *denkend* Subjekt ...” (B155). In this sentence, “I” is regarded as intelligence and thinking subject, meaning that intelligence is thinking subject.

Then, what is this “I”? Kant explains Descartes’ “I think, therefore, I am (*cogito, ergo sum*)” perfectly. *Cogito* is better than “I think”, because in *cogito* there is no subject at all; it expresses thinking activity itself. Kant explains *cogito* in detail in the part of *CPR* on rational psychology (A346/B404 f) and adds relevant paragraphs in the categorial deduction in B edition (B132 [§16]); he confirms that “*sum*” (I am) comes from the self consciousness of “*cogito*”. Consciousness is an important term.

---

3 Kant say “this spontaneity that I call myself [is] an intelligence” (B157–8).

*Cogito* is the activity of *Gemüt*. To say it precisely, it is the activity of one kind of power of *Gemüt*. Kant explains this in two ways. One is to say it is the activity of the synthetic unity of pure apperception (B132 [§16]); the other is to say it is the power of determination or the spontaneity of understanding (B157 [§24]). Kant finds an important thought, that the activity is conscious of itself. Kant uses self-consciousness to refer to pure apperception and to explicate the thinking subject or "I". In particular, through the leap from *cogito* to *sum*, according to Kant, self-consciousness is transformed into the I as intelligence. The reason this Intelligence cannot be regarded as existing is that human beings do not have the capacity of intellectual intuition

Here what is most puzzling is how self-consciousness is transformed into intelligence. To see it literally: if the "I am" exists, the "I" is established; however the "I" here cannot be understood as a sensible being, but as an intelligence. The formal deduction lies in the inference about ideas in Transcendental Dialectic. The concept of the absolute unity of the thinking subject is the synthetic use of categorial inference by reason (A335/B392). In short, reason transforms self-consciousness into intelligence (noumenon) through the activity of inference.

### 3. Analysis of Consciousness and Intellectual Intuition

Because intellectual intuition has been used repeatedly by Mou Zongsan in contemporary Chinese philosophy, it has become a vital term, so it is necessary to clarify its meanings here. Mou Zongsan was unsatisfied with Kant's doctrine of intellectual intuition; he thought it wrong for Kant to deny that human beings have intellectual intuition. The achievements of Chinese philosophy have proved that human beings have intellectual intuition.

Intellectual intuition is a term Kant has come up with when he discusses whether the entities of understanding (*Verstandeswesen*) can be used positively. Understanding (or its spontaneous activity) stipulates sensible intuition through categories and constructs the concept of experience; this means constructing the sensible being from the point of view of ontology. If entities of understanding exist, then it should happen that they spontaneously determine intellectual intuition through using categories. However, human beings have no power of intellectual intuition; therefore, it is impossible for entities of understanding to be determined. This means intellectual intuition would be correlated

with the understanding (spontaneity), and it would be a necessary premise for understanding to determine it.

However, the *zhide zhijue* (智的直覺, intellectual intuition) that Mou Zongsan has drawn from Kant's discussion is a higher cognitive power, whose function is to intuit *liangzhi* (良知, innate knowledge), namely *tianli* (天理, heavenly principle: i. e., 仁, *ren*; 義, *yi*; 禮, *li*; 智, *zhi*), has little to do with constructing entities of understanding. Obviously, the *zhide zhijue* Mou has introduced does not parallel Kant's intellectual intuition. Mou Zongsan has also argued that Kant claims freedom is a postulation. This refers to his point of view that *liangzhi* (corresponding to Kant's practical reason: i. e., autonomy, freedom) is not revelation. Perhaps Mou Zongsan's interpretation is not exact enough. If we translate Mou's *zhijue* (直覺) as consciousness (*Bewußt*) instead of intuition, the two problems above both are solvable. The meaning of consciousness includes awareness and perception; as such, it can be understood as meaning "intuit", "watch", "notice", etc. Kant used nearly all these connotation of it in his works. Sense means judge sometimes. Similarly, a connotation of being conscious includes judging. Its meaning includes sensible (aesthetical) judging and intellectual judging. *Cogito* is an activity of understanding, the activity of spontaneity in determining the manifold of sensible intuition. Self-consciousness is the activity of *cogito* reflecting back to think itself. The spontaneity becomes conscious of itself. Here "to become conscious of" means to view. Thus self-consciousness is spontaneity (pure apperception) viewing itself. The activity of pure apperception to determine the manifold of sensible intuition is sensible consciousness. The activity of pure apperception (*cogito*) viewing itself is intellectual consciousness. Self-consciousness is intellectual consciousness. What parallels Mou Zongsan's *zhide zhijue* is not intellectual intuition, but intellectual consciousness, self-consciousness (i. e., the activity of spontaneity—*cogito*, pure apperception—viewing itself).<sup>4</sup>

---

4 In *CPR* Kant writes (Bxl): "Wenn ich mit dem *intellektuellen Bewußtsein* meines Daseins, in der Vorstellung *Ich bin*, welche alle meine Urteile und Verstandeshandlungen begleitet, zugleich eine Bestimmung meines Daseins durch *intellektuelle Anschauung* verbinden könnte, so wäre zu derselben das Bewußtsein eines Verhältnisses zu etwas außer mir nicht notwendig gehörig." In the translation of W. S. Pluhar (Indianapolis: Hackett, 1996): "[This empirical consciousness of my existence contrasts with] the *intellectual consciousness* of my existence that I have in the conception *I am*, which accompanies all my judgments and acts of understanding: if with that intellectual consciousness of my existence I could at the same time link a determination of my existence through

What corresponds to the Chinese word *guan-zhao* (觀照) in Kant's philosophy is *Bewußtsein* (consciousness), not *Anschauung* (intuition). Considering self-consciousness is the self-viewing of spontaneity (i. e., *cogito*), revealed to itself, the problem that liangzhi (良知, practical reason, freedom, autonomy) is a revelation and not a postulate has been solved.<sup>5</sup> Until now, I still have much doubt about the discussion of Kant's neglecting to interpret and talk about *Bewußt(sein)*. Maybe this is a vital step that can explain how *cogito* is transformed into “*sum*” or “*I*”, all at once, without resorting to the inference of reason.<sup>6</sup>

It is necessary to add that *cogito* should not be divided into I for one part and think for the other. The word *cogito* is wonderful because it directly represents the activity of spontaneity itself. On my own account: *cogito* → *sum* → I. To express it in English, this means that “*I*” equals “*I* think”. Therefore, I am the representation that is formed by the self viewed as “thinking”. It is not “*I* am thinking”. There is no “*I*” existing. There is “thinking” (*cogito*) thinking. “Thinking” reflects on itself and regards itself as “*I*”. In German, “denke” contains “*ich*”; this is not correct grammatically. As a result, this “*ich*” has to be uttered. In English, merely to say “think” does not imply “*I*”; it is necessary to say “*I*” explicitly. I am deeply confused when reading the following expression: “the I of I think”, even “the I of cogito”. Therefore I think if Kant had noticed conscious(ness) more and had further interpreted it, and also had given it a more important role in his philosophy, we would have apprehended his meaning more profoundly and more accurately.<sup>7</sup>

---

*intellectual intuition*, then this determination would not include necessarily the consciousness of a relation to something outside me.”

- 5 Kant says of freedom (5:4, tr. W. S. Pluhar): “this idea reveals itself through the moral law.”
- 6 The two examples below, from *CPrR*, obviously show the significance of the concept: “ebendasselbe Subjekt, das sich andererseits auch seiner, als Dinges an sich selbst, bewußt ist” / “the same subject, who on the other hand is also conscious of himself as a thing in itself” (5:97, tr. Pluhar); and “das *Sinnenleben* hat in Ansehung des *intelligibelen* Bewußtseins seines Daseins (der Freiheit) absolute Einheit eines Phänomens” / “the *life of sense* has in regard to the *intelligible* consciousness of one's existence (the consciousness of freedom) [the] absolute unity of a phenomenon” (5:99).
- 7 Kant uses the word *Bewußt* repeatedly, although he discusses it only occasionally. That there are different explanations about its meaning shows that he has not yet put it forward as an important thesis. Take as an example that in the first edition the title refers to the thesis that the soul is a person is: “Was sich der numerischen Identität seiner Selbst in verschiedenen Zeiten bewußt ist, ist so-



#### 4. Intelligence as the Moral Subject

The origin of the thinking subject has been clarified. Furthermore, “I” or intelligence is the moral subject. In studying practical reason, Kant claims intelligence arises from the concept of causality.<sup>8</sup> Naming the cause of moral action as will, concerning the practical use of reason, there is no disagreement either among philosophers from different factions, or between common people. Kant’s contribution lies in his distinguishing of free will—that is, eliciting freedom from the concept of the will. The argument Kant has given attempts to prove that concepts of the free will, autonomy, and self-legislating are alternating concepts and meanwhile elevate our mental power to entity-intelligence.

In the Introduction to *Critique of Practical Reason*, Kant declares that this *Critique* will show that pure practical reason exists. In other words, human beings are free. Especially important is that Kant declares that practical reason itself, without any collusion with the speculative, provides reality to a supersensible object of the category of causality: freedom (5:6). Here, the appositive of freedom is causality, not supersensible objects.<sup>9</sup> What concerns Kant is whether the subject of moral behavior has objective reality.

The aim of my essay is to answer the following question: what is the human being? Here we focus on the question: what is the moral subject? In *GMM* Kant uses intelligence several times. He says a rational being must itself be regarded as intelligence. He argues in general here that human beings have one kind of power, the power of reason, that can distinguish us from everything else, even from ourselves as affected by objects. Reason is pure spontaneity (*Selbsttätigkeit*) and is higher than understanding (also spontaneity). Understanding has to be ap-

---

fern eine Person.” Through the whole discussion in this paragraph, Kant understands the word *Bewußt* in a sense that amounts to the activity of “intuiting” (A361–6). But in the deduction of categories discussed in the second edition, this word is used many times in a variety of senses. Only at the end of section 16 does the meaning of the activity of “intuiting” appear (B135): “Ich bin mir also des identischen Selbst bewußt; ich mir einer notwendigen Synthesis derselben a priori bewußt bin.”

8 In *GMM* (4:453) Kant writes: “Every rational being reckons himself qua intelligence as belonging to the world of understanding, and it is simply as an efficient cause belonging to that world that he calls his causality a will.”

9 Kant writes (5:132): “Diese Postulate sind die der *Unsterblichkeit*, der *Freiheit*, positiv betrachtet, (als der Kausalität eines Wesens, so fern es zur intelligibelen Welt gehört,) und des *Daseins Gottes*.”

plied to sensation, for its duty is to make the sensational representation belong to rules; however, reason is pure, so its duty is to give limits to understanding, to distinguish the world of sensation and the world of understanding (4:452).

However, as I see it, Kant has not made clear how the intelligence is formed. His argument is to show that the thing in itself is a member of the supersensible world, and this is just the world of understanding; therefore, intelligence is just the thing in itself. Intelligence is what speculative reason produces, and is only here does it acquire its moral significance. Moreover, the doctrine referred to earlier, that the concept of the thing in itself is created by the power of feeling or judgment, looms large, then disappears.

The thing in itself had been a thought, but here intelligence has to have objective reality. Why? I think the answer is rather simple: it is because of the needs of real living. As thinking beings, it matters little that our intelligence has no objective reality; as moral subjects, however, intelligence must be responsible for our wrong doings, so its objective reality is required. To put it another way, in real life, human beings have been made responsible for their wrongdoings; now what philosophy does is to construct theories to interpret the grounds for these social phenomena.

The opinion of common people is that everyone's actions are autonomous; therefore, we can suppose there is a moral subject in every human being, dominating his or her own behavior. However the real case is that there is not such a subject at all. What dominates one's behavior is the free will. The free will is one part of our mental powers. Power is not an entity. Yet the opinion of common people is to transform the power into being. Kant's mission in his moral philosophy is to construct theories for elevating free will to the moral subject. The possible logical mistake in Kant's argument is that he equates the thing in itself with the world of understanding.

The remaining problem is: the requirement for the objective reality of the moral subject originates in the fact that human beings have to be responsible for their own wrongdoings and should deserve the penalties for their wrongdoings. My question is this: should we discuss this problem in the philosophy of right?

## 5. Is the Criminal Subject a Phenomenon or Intelligence?

As is well known, human beings are fallible, capable of committing crimes. The question now arises: which side of the human being implements such acts? Kant emphasizes that all kinds of being must be united. We must not only recognize the duality of human beings, but also regard them as necessarily integrated into one subject. He writes (4:456): “die Kausalität nach dem Gesetze der *Naturnotwendigkeit*, *bloß der Erscheinung*, die *Freiheit* aber eben *demselben Wesen*, als *Dinge an sich selbst*, beizulegen.” He also stipulates that the unity of pure speculative reason and pure practical reason is based on reason itself, so the unity is necessary (5:95); practical reason takes priority (5:121).

Kant also takes seriously the problem of the worthiness to be happy in proportion to one’s morality. He spends much energy expounding it. For me, however, the most relevant application is the punishment for criminal activity that he refers to but does not demonstrate in much detail. Kant’s discussion basically centers around the dual structure of the human being. Much of his difficulty results from this dual structure. We will suggest a new structure to make it more coherent.

The concept of the thing in itself is discussed in the first *Critique*. What corresponds to our empirical knowledge is the phenomenon, not the thing in itself that causes the phenomenon. Provided that Kant has postulated that sensation is caused by the affection of objects and that sensibility is perceptivity, we must postulate the thing in itself as the source of sensation.<sup>10</sup> This is correct when Kant says the thing in itself is a thought. Substance is a category of thought, and the thing in itself is a thought as well. As we have mentioned earlier, the thing in itself is caused by the assertion of the judgment of feeling.

Grande dame interrogates Fichte as following: is your wife the non-I that is created by your Self? The one whom he hugs is certainly not the creation of his self. However, women, wives, beauties of nature, etc. are all phenomena, and thus all creations. The world is a great flux; there is nothing for us to watch. Things are made by watching. The wife of Fichte should be regarded as such. There is nothing at all in the world; what is there is only flux. All aspects of physical relations, moral behavior, relations of right, deliberations of beauty and affirmations of property are merely things that are all created by the mind.

---

10 “Die Fähigkeit (Rezeptivität), Vorstellungen durch die Art, wie wir von Gegenständen affiziert werden, zu bekommen, heißt *Sinnlichkeit*” (A19/B33).

It is obviously wrong to take the thing in itself as referring to flux.<sup>11</sup> Therefore, it is correct to say the thing in itself is a thought. However, what is it that causes sensation? It is obviously not something, not things (neither a phenomenon nor the thing in itself) at all, but flux. We cannot say sensation is caused by the affection of the thing in itself on sense organs. What causes sensation is not anything at all. Therefore, in the first *Critique*, the thing in itself is merely a mediate, transient concept. To prove all things belong to the sphere of phenomena, it is necessary to resort to the concept of the thing in itself. Once the proposition that all things are only phenomena is established, it is necessary to reject the concept of the thing in itself.

After Kant proposes the concept of intelligence, he thinks it also as the thing in itself. As things created by the mind, they are quite similar. However, when first put forward, the thing in itself is assumed as the source of sensation. But the intelligence proposed when discussing pure apperception and self-consciousness must not be regarded as the source of sensation.

When we enter into the world of morality, it is unavoidable to elicit the concepts of free will, the intelligent world, morality and happiness, crime and penalty. Here looms a deep dilemma: free will is pure practical reason, the intelligent world is a pure world, and crime and penalty are not in this world. As a result, these burdens have to be imposed on the sphere of nature. Then what is the sphere of nature? It is the sensible world, the world where phenomena reside. Only empirical knowledge and concepts exist in this world, where the wrongdoings and the crimes human beings commit are described and communicated. However, human beings will not commit crimes and suffer wrongdoings in this world.

Where can human beings commit wrongdoings? Is the subject committing crimes a phenomenon or intelligence? It is certain that the intelligence cannot commit errors at all; therefore, Kant has to attribute wrongdoings to phenomena, yet this is obviously implausible. It is only possible that wrongdoings are in the great flux. Errors and crimes all belong to the judgment that is about the facts observed and is made according to morals and laws. The facts observed belong to what Kant called the sphere of nature, that is, the world of knowledge and entities

---

11 It is interesting that Kant is clearly aware of this (A364): "the dictum of certain ancient schools, that everything in the world is *in a flux* and nothing is *permanent* and abiding, cannot be reconciled with the admission of substances."

of sensation. In this world we can describe the lives of human beings, we can transmit the living of human beings, but there is no real living at all. The world of intelligence is also a world of conception that can restrict the behavior of human beings to the scope that rules allow. It is wrong for Kant to call this world “archetypal”. It can be called the world of models, but not archetypal. The archetypal world is the world of things in themselves. The world of phenomenon and the world of noumenon are both the images of it. However, things in themselves cannot constitute a world at all, for there is only the great flux.

As a result, Kant should enlarge human nature from twofoldness to threefoldness. It is not possible that there are no concepts of substance and person. All the relationships concerning property and ethics are based on these two concepts. Substance and nature, person and freedom are all concepts produced by the mind. The spheres of nature and freedom are both worlds produced by the mind. Worlds are concepts and sentences produced by the mind. The real is the great flux that is unspeakable. Therefore, the greatest difficulty in Kant’s philosophy lies in his doctrine of the thing in itself.

## 46. Mou Zongsan's Interpretation of the Kantian *Summum Bonum* in Relation to Perfect Teaching (*Yuanjiao* 圓教)

Annie Boisclair

The work of New-confucianist Mou Zongsan (1909–1995) has been strongly influenced by the writings of Kant. Mou uses Kant's classic theories to support his own reinterpretation of Chinese philosophy, offering new and exciting perspectives on a major contemporary problem of Chinese culture: instead of viewing oriental and occidental philosophy as isolated concepts that are mutually exclusive, Mou proposes that Eastern thought and Western ideas should fecundate and stimulate each other. Moreover, as he regards Kant's approach to morality as compatible with Confucian views, Mou considers the Kantian philosophical system as the closest Western equivalent to Confucianism.

One of the most interesting topics Mou investigated in his later years is the *summum bonum*. This concept is originally derived from the ancient Greeks; Mou, however, who also refers to it as the "Perfect Good" (*yuanshan* 圓善),<sup>1</sup> bases his work mostly on Kant's version, the harmony between virtue and happiness guaranteed by God, as described in his *Critique of Practical Reason*. According to Mou, this concept is the pinnacle of Western thought just as the Perfect Teaching represents the peak of the Chinese culture. Moreover, in his opinion, it is possible to further clarify and elaborate Kant's theory through the Perfect Teaching. He firmly believes these are essential concepts and require extensive debate, since the relation between virtue and happiness should be universal. The Perfect Good has a strong appeal to Mou, particularly because he is himself a Confucian and morality is the nexus of Confucian tradition. While this is true also for Christianity, Mou argues that few, if any, human beings would be willing to make sacrifices in order to achieve virtue without also being rewarded with happiness. Bliss is desired

---

1 See the reason for his translation in Zongsan Mou, 中國哲學十九講 (*Zhongguo zhaxue shijiu jiang; Nineteen discussions on Chinese Philosophy*) (Taipei: Xuesheng Shuju, 1983), 329. All translations from Mou's writings are my own.

by everybody. In contrast, Kant thinks those two elements of the *summum bonum* go together (5:110):

That virtue (as worthiness to be happy) is the *supreme condition* of whatever can even seem to us desirable and hence of all our pursuit of happiness and that it is therefore the *supreme* good has been proved in the *Analytic*. But it is not yet, on that account, the whole and complete good as the object of the faculty of desire of rational finite beings; for this, *happiness* is also required, and that not merely in the partial eyes of a person who makes himself an end but even in the judgment of an impartial reason, which regards a person in the world generally as an end in itself.

Virtue has a privileged status: it is the supreme condition that allows people to consider their acts, words or objects as desirable. Nevertheless, by itself, it does not represent the Perfect Good, because the latter comes into existence only when virtue is linked with happiness. Because of this, their relationship is essential both for persons in search of happiness as well as for those pursuing virtue above all. By contrasting the Stoic, Epicurean and Kantian interpretations of the *summum bonum* with Mou's interpretations and modifications of this concept, we aim to distill the conceptual advances Mou added to the theories of the Perfect Good and come to the surprising conclusion that he, by utilizing Buddhist thoughts to improve a Western school of thought, attempts to further the cause of Confucianism.

## 1. Perfect Teaching: A Buddhist Concept Adopted by a New-Confucian

In China, Buddhism has evolved into a multitude of schools, each with their particular doctrines and practices. The concept of Perfect Teaching has been created by the Tiantai<sup>2</sup> and the Huayen<sup>3</sup> schools. Despite

---

2 Tiantai is a school founded by the Chinese Zhi Yi (538–97). It synthesizes the idea of other schools, giving a place to all sūtras, and judges the Hinayana (Lesser Vehicle) and the Mahayana (Great Vehicle) as part of the Buddha's authentic teaching. In fact, Zhi Yi created a wide syncretic system that gives an overview on Buddhism while it finds a place for each practice and doctrine. His classification of doctrine categorizes all Buddhist teaching as an expression of Truth aimed at specific circumstances and students at heterogeneous levels.

3 The Huayen school was officially established by Fazhang (643–712); nevertheless, the true beginning of the school is dated around 581. The Huayen doctrine perceives itself as perfect, since it encompasses the totality of phenomena. All

the fact that these two schools differ in several aspects of their classifications of doctrine (*panjiao* 判教),<sup>4</sup> both regard the Perfect Teaching as the noblest doctrine and as Buddha's most perfect and complete teaching. As Mou esteems the Tiantai classification as the more systematic one, he uses it to clarify Kant's theory of the *summum bonum*. The Chinese word for Perfect Teaching is *yuanjiao*, or 圓教. The first character, 圓 represents an enclosure confining a circle; this conceptually refers to the idea of complete perfection encompassing everything without introducing division; it unites rather than separates. The second character, 教, means teaching or doctrine. Another translation of 圓教 sometimes found in the English literature is "Round Teaching", but Mou disqualifies this term as not representative of the true meaning of the concept of *yuanjiao*.<sup>5</sup> Unfortunately, neither Mou nor his students provide a clear definition of Perfect Teaching. However, Paul L. Swanson's illustration appears to correlate well to Mou's writings:

The Perfect teaching is truth as it is. It is perfectly complete; the truth, the whole truth, and nothing but the truth. It is subtle, inconceivable, beyond verbalization and conceptualization. It is reality as perceived perfectly by the Buddha. It is the Middle Path, which is the insight into reality as simultaneously lacking in substantial Being yet conventionally existent.<sup>6</sup>

This teaching is perfect because it depicts the world as it really is. It accounts for all dharmas,<sup>7</sup> that is for all phenomena. It accounts for the existence of both pure and impure dharmas; this is important, as individuals have to deal with and process both in order to achieve Buddhahood. This holistic view is one major reason Mou favors the Perfect Teaching over many other Buddhist doctrines that according to him either unilaterally negate the existence of pure or impure dharmas, or

---

phenomena become One and One is divided into multitude. Therefore, all phenomena are dependent on each other and mutually conditioned.

- 4 Since only the Tiantai classification of doctrine is relevant here, we will not describe the Huayen classification. The Tiantai divided the teaching of the Buddha into eight doctrines revealed in five periods (*wushi bajiao* 五時八教). The eight doctrines are even more specifically divided in two further groups: the four means of transmission (*huayi* 化儀) and the four degrees of profoundness (*huaifa* 化法). The Perfect Teaching is the highest degree of profoundness.
- 5 See Mou's argumentation in *Nineteen Discussions*, 322–4.
- 6 Paul L. Swanson, *Foundation of T'ien-t'ai Philosophy: The Flowering of the Two Truths Theory in Chinese Buddhism* (Berkeley: Asian Humanities Press, 1989), 11.
- 7 In the context of this text, dharmas mean: all things, all phenomena as they appear to sentient beings.



allow for their coexistence only based on an argumentation that contains major flaws.<sup>8</sup>

The Perfect Teaching itself, as well as the complexity of the reality it describes, is beyond conceptualization. Mere words are not sufficient to characterize it; they would only rigidify and imprison the ever moving, fluid truth. Hence, Perfect Teaching uses a non-disputable language that breaches canonical logic; Mou qualifies this as a non-analytical discourse that also encompasses all reality. Perfect Teaching does not render judgment; it only recognizes purpose that is appropriate to certain situations or a given level of understanding. Sentient beings can reach enlightenment because the world is as it is, with its beauty and its horror. It is not valid to judge the universe as good or evil, because everything a being encounters or does, including the pure and the impure, is a necessary step on its way to attain nirvana. As all sentient beings have a Buddha nature within them, they all have the possibility to reach enlightenment and become Buddha. This reflects the Middle Path of the Tiantai, where one views life and everything surrounding one as an illusion, yet perceives reality as a necessary prerequisite on the path to Buddhahood. Perfect Teaching encompasses all this; therefore, it is complete and perfect.

## 2. Two Unbalanced Versions of the *Summum Bonum*

Both Mou and Kant refer to the Stoics and the Epicureans who have developed theories with virtue and happiness playing a fundamental role. They mostly critique the arguments and conclusions of both schools. In regard to the *summum bonum*, the Greek schools do not treat virtue and happiness in an equilibrium, as favored by Kant and Mou, but rather place high importance on one of them while neglecting the other. Moreover, both Stoics and Epicureans derive their view of *summum bonum* analytically, while Kant and Mou argue that happiness and virtue have a synthetic relation.

The Stoics rank virtue above all and claim that all happiness emerges from it. As such, virtue is the sole condition that allows a being to reach a blissful life. The accomplishment of a virtuous act is seen as inevitably

---

8 According to Mou, this is for instance the case with the Mere Ideation school (*weishi* 唯識) that offers very distorted arguments to allow the realization of pure dharma.

and most profoundly fulfilling, more so than any other good. Thusly, virtue does not need to be prompted by any other kind of reward. Kant wrote (5:112): "The Stoic maintained that virtue is the *whole highest good*, and happiness only the consciousness of this possession as belonging to the state of the subject." Therefore, Stoic theory as interpreted by Kant proposes that the consciousness of performing a virtuous act, of being a moral being, is equivalent to happiness and is entirely sufficient to achieve Perfect Good.

Stoics do not account for happiness; it is only regarded as valuable as an associated subordinate to virtue. Mou agrees with Kant's Critical view: "In this way, only the aspect of virtue is realized, the aspect of happiness was diverted; it does not have an independent meaning. The highest good must include the two aspects of virtue and happiness; they both have a relation of subordination, but one cannot cancel out the other."<sup>9</sup> Since only virtue is important and the sole key to happiness for the Stoics, the latter is only an appendix without any value of its own and contained within virtue. The link between virtue and happiness might explain why some individuals feel inclined toward virtuous acts. However, the Stoics fail to acknowledge that this does not mean that happiness cannot exist outside virtue.

For the Epicureans the word "happiness" relates to an entirely different reality. In fact, they often use "pleasure" rather than "happiness" to refer to the enjoyable sensation emanating from the experience of the physical and mental realms. In Epicurean theory, happiness is more important than virtue; moreover, it is not related to moral and transcendental aspects,<sup>10</sup> but mostly to physical and psychic perception. For Epicurus, absence of pain paired with mental and physical stability are key requirements of happiness. Epicureans search for pleasure and its prerequisites to reach the *summum bonum*. According to Epicurus, nature provides everything human beings need to enjoy life and uses pleasure as an indicator showing human beings that they are on the path of the Perfect Good. In contrast to the repulsive suffering sentient beings try to avoid, pleasure is a joy and attractive, and therefore a good datum on a beneficial path.<sup>11</sup>

---

9 Mou, 373.

10 Geneviève Rodis-Lewis, *Epicure et son École* (Paris: Gallimard, 1993), 215.

11 A. A. Long, *Hellenistic Philosophy: Stoics, Epicureans, Sceptics* (London: Duckworth, 1986), 62.

Nevertheless, Epicureans do not preach that all pleasures should be pursued at all cost, but also counsel moderation. Ideally, one should reach an equilibrium characterized by the absence of pain, presence of pleasure and avoidance of overindulgence. In the physical realm, bodies signal their basic desires and needs often through unpleasant sensations; humans can usually address these matters simply, for instance through the intake of food. However, once the stomach is full, one should stop eating. Pleasure beyond need can bring great suffering, particularly when sought out of greed or lasciviousness. Furthermore, unchecked pleasure and luxury endanger the stability of mind and body: the fear of losing privileges and possessions may well outweigh the pleasures that derive of them. Moreover, one should be able to suffer a minor pain in order to achieve a great pleasure or to avoid minor pleasure in order to prevent greater suffering. According to Epicurean theory, a major challenge on the way to the Perfect Good is to distinguish between necessary and natural versus superfluous and dangerous desires, and to achieve a balance of appetites and satisfactions as a prerequisite to a peaceful and stable bliss of body and mind.

Virtues serve to guide an individual on this difficult route. Epicurus writes about the four classical Greek virtues: prudence, justice, moderation, and courage. Epicureans do not regard them as an end, but merely as important means that allow human beings to reach happiness. The virtues are attached to bliss, because they are necessary to attain it, and therefore regarded as useful tools that help to organize desires and their satisfaction, and to check unhealthy desires and fear. This view of virtue is not founded on moralism; while having a moral connotation, it is not essentially moral. Without pleasure, virtue is nothing; it is without power. In contrast, bliss is a hint that allows humans to achieve harmony with their surrounding, in the midst of their needs, and with what is available. The Epicurean's happiness seems simple and natural: people have appetites they must fulfill to ensure health of the body and peace of mind; with the help of virtues, they can judge the quality of their desires and act only on those that are necessary to keep a state of happiness.

In Mou's opinion the Epicurean view is unbalanced in the opposite direction of the Stoics; they place too much weight on happiness while they underestimate the contribution of virtue to bliss by regarding it merely as a means rather than an end. This raises substantial problems, because not everything that pertains to bliss is necessarily compatible with virtue. For example, biting a fresh crisp apple of the orchard

next door can be a real joy, but does not entail any moral issues—if the apple was obtained legitimately and produced ecologically. Because Mou postulates that the Perfect Good necessitates a synthetic relationship between virtue and happiness, with these two being balanced in favorable equilibrium, he does not regard the Epicurean *summum bonum* as appropriate. However, his interpretation of the Epicurean school of thought remains partially controversial. In his view, the Epicureans transformed happiness into a moral principle,<sup>12</sup> but in Epicurean theory virtue is neither based on moral considerations nor does it exist to safeguard morality.

### 3. Kantian *Summum Bonum*

In contrast to the Greek schools discussed above, Kant does not regard virtue and happiness to be in a state of unilateral dominance, where either one may subjugate the other principle to deprive it of its independent fundamental value. Kant's Perfect Good is characterized by a synthetic harmony, with virtue and happiness resting in a favorable equilibrium (5:110). Perfect Good becomes a whole—complete, absolute, and perfect. Mou regards Kant's *summum bonum* not as an analytic<sup>13</sup> relation reminiscent of the Stoic or Epicurean theorems, but rather as a synthetic one. Nevertheless, even in this synthetic classification, this relationship is not exactly egalitarian. Happiness can exist independent of virtue; however, on its own it does not lead to the *summum bonum*. In contrast, even virtue initially devoid of happiness might be sufficient to lead to the Perfect Good. This is because happiness arises from virtue, as granted by God; therefore virtue remains intimately but not necessarily temporally linked to happiness. Happiness then remains essential in the accomplishment of the Perfect Good because it is a necessary consequence of any virtuous deed.

Perfect Good arises through a synthetic relation (5:112–3) of two elements of very different natures. Because of their difference, one cannot understand the *summum bonum* analytically; neither is it possible, by experience or relying on empirical data, to discover that those two concepts are intimately linked. Moreover, a separate analysis of the essence of either virtue or happiness cannot unravel a link between them, be-

---

<sup>12</sup> Mou, 373.

<sup>13</sup> Mou, 373.

cause analytically neither relies on the other. One can experience forms of happiness while completely ignoring underlying moral issues; conversely, some of the most virtuous acts may be painful enough to obliterate any feeling of wellbeing. Therefore, the comprehension of this synthetic relation between happiness and virtue can only be transcendental. In other words, this interconnection belongs to the suprasensible, despite the fact that the actions toward Perfect Good are anchored in the sensible world (5:119). Regardless, their combination is a practical necessity and Mou explains this necessity by the fact that virtue without happiness would be too tragic to be appealing to humanity.<sup>14</sup>

According to Mou, the Kantian version of the *summum bonum* represents progress for philosophy and is a remarkable accomplishment. Furthermore, he considers Kant's conceptual analysis of virtue and happiness as very coherent. However, he deplors that the theory of the Perfect Good is neither as well developed nor as clear as Kant's theories on the moral law or the categorical imperative.<sup>15</sup> Mou agrees that happiness and virtue synergize in the *summum bonum* if they are in an appropriate equilibrium, with moral values nevertheless remaining paramount over felicity, and that this relation of reciprocal dependency is synthetic. Even so, Mou goes further than Kant by arguing that this relation is synthetic only from a human point of view. In human beings' everyday life, happiness may exist without necessitating morality, as well as virtuous action that does not lead to bliss. Because of this, Mou concludes that an analytic standpoint does not necessarily reveal the existence of a relation between those two elements. Consequently, the link between them must be synthetic, and this requires the guaranty that those parts will stay joined. Kant postulates that this guaranty is given by God, who is the omniscient creator of the world; therefore, Mou concludes that God perceives the relation between virtue and happiness as analytic.<sup>16</sup> The Almighty possesses intellectual intuition; consequently, He has the knowledge of both the empirical and transcendental world, including the relation between the two elements of the Perfect Good as inseparable, thanks to His will. Only in His eyes is this an analytic relation; mere mortals perceive the relation only synthetically.

---

14 Mou, 377.

15 Mou, 378.

16 Mou, 382.

#### 4. Deepening the *Summum Bonum* through the Perfect Teaching

Despite agreeing with many of his thoughts, Mou also finds a major flaw in Kant's description of the *summum bonum*; specifically, its strict dependency on an outside power attributed to God. In other words, Kant postulates that God's power is absolutely necessary to ensure the existence of the Perfect Good in this world. Mou proposes to solve this problem by clarifying and improving Kant's theory using elements of the Perfect Teaching. While virtue is generally associated with the mind, Mou insists that happiness is in essence linked to the world of phenomena; the possibility of experiencing felicity depends entirely on the existence of dharmas that are therefore essential to ensure the access of human beings to happiness in this world: "The idea of happiness rests on the existence of dharmas. The existence of dharmas is precisely the existence of the actual world."<sup>17</sup> Therefore, the material realm is a prerequisite for the existence of happiness. Mou argues that Perfect Teaching is guarantying the existence of all dharmas and their necessity, unlike other Buddhist schools. Kant attributes the creation of the universe to God, who guarantees the existence of all phenomena. God also possesses intellectual intuition, and He furthermore represents absolute perfection. Therefore, Mou compares God to the Perfect Teaching in Buddhism:<sup>18</sup> "What Perfect Teaching represents is the level of God, it is the absolute. In this case, the necessity of the existence of dharmas is entirely preserved."<sup>19</sup> Thereby, Mou equates God, creator of the universe and guarantor of the existence of all phenomena, to the doctrine of Perfect Teaching that advocates the necessity of the existence of all dharmas. This comparison seems somewhat inadequate, since the two concepts are different in terms of form and purpose. He probably means that God represents a perfect system, just as the Perfect Teaching does, both encompassing and ensuring the existence of every single thing. Therefore, by making use of the theory of Perfect Teaching, where the Buddha guarantees the existence of all dharmas, no concept of God is required to fulfill this function.

Mou criticizes Kant's attempt to use God as a means of ensuring the realization of the *summum bonum* for a second reason. He argues that, if

---

17 Mou, 378.

18 Mou, 379.

19 Mou, 382.

God alone creates harmony between a person's virtue and its fair amount of happiness, human beings would be devoid of both responsibility and control over their happiness. In this case their only way of influencing their happiness would be through their own virtue; yet even then, the amount of happiness for a virtuous deed would not be self-determined but rather determined by God, who is held responsible to harmonize the two elements. In his reinterpretation of the *summum bonum*, Mou proposes to relieve God from the responsibility of fairly distributing happiness according to virtuous action. Again, he replaces God by the Buddha as described by the Perfect Teaching. Also referred to by Mou as the Perfect Buddha (*yuanfo* 圓佛), he represents the absolute in the Perfect Teaching and acts as guarantor of the existence of all dharmas. From several angles, this permutation between God and Buddha is rather equivocal: there is only one God, but there is more than one Buddha; God created the world, Buddha did not; nobody can transform into God, but everybody can become a Buddha; God is thought of as being external to humans, while the essence of Buddha is present in all sentient beings. Mou emphasizes the fact that God and Buddha are the beings of reference in their respective doctrine, and that both ensure Perfect Good. However, Mou's substitution of the Kantian God by the Perfect Buddha has a tremendous impact on the doctrine of the *summum bonum*, because all humans have the capacity to reach buddhahood. Consequently, happiness is no longer the concern of a supreme being, but becomes the personal responsibility of each human. Being able to influence both virtue and happiness, the relation between those two elements becomes, as for the Kantian God, analytic.<sup>20</sup> However, Mou ignores the question whether human beings have the capacity to measure adequately the quantity and quality of happiness in regard of its virtue. The *summum bonum* requires more than happiness and the control of it; it critically depends on an appropriate equilibrium between virtue and happiness. To aspire to the Perfect Good, one must be worthy of happiness. However, this would seem to entail that one has to achieve enlightenment and become a Buddha in order to be able to perform this daunting task.

---

20 Mou, 383.

5. Serving Confucianism through the *Summum Bonum*

Intriguingly, Mou spices up his work with a surprising twist. Despite having elaborately reinterpreted the Kantian *summum bonum* using elements of the Perfect Teaching, he denies the concept of Perfect Good a rightful place in Buddhism itself. He argues that morality is not an essential concept of Buddhism, yet that: "Only by discussing morality can there be the concept of Perfect Good."<sup>21</sup> Moreover, he suggests that Confucianism provides a better theoretical background to accommodate this concept. According to Mou's interpretation, the core of Confucianism is built around the concept of morality and contains a substantial discussion on the nature of good and virtue. This is related to Kantian theory, where good is established by the moral law, but not determined by an outside object (5:64). In Confucianism, the distinction between good and bad is derived from human nature, itself being neither; our neutral human nature serves as the criterion allowing us to identify the good and the evil. This criterion cannot be judged to be good or evil based on another criterion; it is the ultimate standard permitting classification of everything as good or evil. Therefore, human nature is absolute: "That which is without good and without evil is called the highest good. The highest good is the absolute good."<sup>22</sup> Everything that can be described by those two adjectives is relative, as it contains both good and evil in different proportions.

Confucianism is not constrained to its investigation of morality. Mou furthermore considers it to fit the criteria of the Tiantai classification of doctrine that determines the Perfect Teaching; therefore Confucianism itself rises to the rank of a Perfect Teaching: "To discuss the problem of good and evil from the standpoint of Confucianism, there also is a Perfect Teaching; besides, its method is the same despite the fact that its content is not the same, because Confucianism has a moral philosophical system."<sup>23</sup> Mou qualifies the aspect of morality in Confucianism as a Perfect Teaching. This is particularly obvious during the period of the Song-Ming Dynasty (960–1644). At this time, some philosophers such as Wang Yang-Ming (1472–1528) use a non-disputable language when talking about good and evil. Also, they regard good and evil without judgment and rather acknowledge them as entwined

---

21 Mou, 383.

22 Mou, 384.

23 Mou, 384.



parts of each other that may change their relative proportions depending on the circumstance and the observer. Furthermore, these writers claim that all human beings have the ability to become sages and act according to wisdom-derived moral judgment without failure. All these criteria make Perfect Teaching perfect and complete.

## 6. Summary and Conclusion

The *summum bonum* is a concept that exists since the ancient Greeks. For the Stoics, happiness is where virtue dwells. Virtue is essential and ruling over felicity. In the Epicurean's theory, virtue rests in happiness. Pleasure serves as a gateway to goodness while virtue acts as a tool to help achieve Perfect Good. Kant advanced this concept to a state where neither virtue nor happiness eclipse the other side, but rather form a harmonic equilibrium. God, who as the Almighty creator of the world is the sole possessor of intellectual intuition, guarantees a favorable ratio between virtue and happiness.

Despite judging Kant's theory as the pinnacle of Western philosophy, Mou detects substantial flaws inherent to it. To solve these, he proposes to use some elements of the Buddhist Perfect Teaching to clarify the process of *summum bonum* and assign human beings to its center. In doing so, Mou substantially advances our views on Kant and forms a particularly interesting bridge between one of the most influential philosophers of the West and the intrinsically different approach taken by an important branch of Chinese philosophy. However, despite achieving an interesting synthesis between these two so different schools of thought, he somewhat surprisingly concludes that in respect to understanding the relation between happiness and virtue, Confucianism provides a more valuable approach than Kantian theory, the Perfect Teaching of Buddhism, or even their synthesis. It seems somewhat ironic that this New-confucian writer utilizes a Buddhist concept to emphasize the qualities of Confucianism; it would be intriguing to analyze in depth the motives that moved Mou to take this approach.

## 47. Confucianism and Things-in-themselves (Noumena): Reviewing the Interpretations by Mou Zongsan and Cheng Chung-ying

*Chaehyun Chong*

### 1. Kant and Confucian Philosophy

Why do we compare Kant's and Confucian Philosophy, two philosophies that occurred in two different cultural backgrounds? This is probably due to their common emphasis on absolute morality and their common attention to morality as the essence of human dignity. In other words, since Confucian's moral subjectivity can be compared to Kant's spirit, regarding morality as originating from pure practical reason or moral will, it is natural to compare Confucianism with Kant's philosophy. Mou Zongsan (1909–1995), especially, claims that Kant's philosophy is the only one in the West that can be compared to Confucianism. He, however, while admitting the strengths of Kant's philosophy and emphasizing its similarity with Confucianism, talked about the superiority of Confucian philosophy over Kant's philosophy. The reason he thinks of Confucianism as being superior is the insufficiency or defect of Kant's division between phenomena and noumena. Mou calls Kant's division transcendental and says it is not sufficient. Everyone would agree with his claim that the division between phenomena and noumena is not empirical, but *a priori* or transcendental. Kant himself surely would not deny this. If so, why did Mou say Kant's distinction was insufficient or defective?

According to Mou, for the distinction between phenomena and noumena to be really transcendental, in other words, in order for the distinction to be well-established, we need to know the conception of noumena. Confucianism can provide Kant's philosophy with that secured understanding of noumena. In Mou's perspective, in order for Kant's philosophy to be a more secured moral metaphysics, the concept of noumena should be non-factual, but value-oriented and so appre-

hended by intellectual intuition without being taken as merely a postulate of practical reason.

Cheng Chung-ying also pays attention to the theoretical similarity between Kant's philosophy and Confucian philosophy. He also claims that Confucianism acknowledges the possibility of knowing noumena while Kant's philosophy cannot. As the method of knowing them, however, he emphasizes not intellectual intuition but cognition through a dynamic process based on his onto-hermeneutical interpretation of Confucianism. The ultimateness that Kant and Mou found in the noumena is provided by his concept of mankind's unending process of interpretation, to borrow his words, the creative creativity. In this paper, I will examine and evaluate the main points of Cheng and Mou in comparing Confucianism and Kant's philosophy.

## 2. The Conception of Things-in-themselves in Kant

### 2.1 Things-in-themselves in Theoretical Reason

Things-in-themselves in theoretical reason have a passive and a limited sense. According to Kant's so-called Copernican revolution, our cognition is not about things-in-themselves, but about objects constructed by our sensible intuition, pure forms of intuition and the categories of the understanding. Things-in-themselves are conceived as corresponding to "things other than phenomena" (A251). In other words, they may be the causes of phenomena that can be thought through the capacity of understanding. Because of this definition of the concept, we cannot have cognition of things-in-themselves as objects. This is because an object is the composition of intuition and concept, and it is not possible for us to have either sensible intuition or pure intuition of things-in-themselves. Things-in-themselves, in this regard, cannot but be a formal or a logical concept that lacks objectivity; it is therefore a void concept. The passivity of this concept of things-in-themselves in theoretical reason, however, does not have only this negative meaning. Rather, through this aspect of the concept of things-in-themselves, we see that the concepts or the categories of our understanding must be used empirically, not transcendently.

## 2.2 Things-in-themselves in Practical Reason

Things-in-themselves in practical reason refer to unlimited transcendental ideas such as freedom, soul, and God; all of these are not related to individual things as they are in theoretical reason. Since things-in-themselves originally meant things thought only through the capacity of pure understanding, they are used in a regulative way in theoretical reason; but these ideas of transcendental reason are revived positively as the postulates—that is, they are needed for moral laws to be effective. In other words, they can never be known or cognized in the framework of theoretical reason because of their unlimitedness. Their existence can only be justified practically in the dimension of practical reason. Of course, this practical justification of their existence just shows that the ideas of practical reason as subjective principles do have only subjective necessity. Human beings, who have a limit as phenomenal beings in the dimension of theoretical reason, disclose their aspect of transcending phenomena in the dimension of practical reason, as autonomous beings who themselves make laws and follow them by themselves. The conception of things-in-themselves that is negative and limited in theoretical reason becomes positive in practical reason.

## 3. Mou Zongsan's Criticism against Kant

Mou Zongsan says Kant's distinction between phenomena and noumena is not well-established. In order for that distinction to hold, Mou believes the concept of things-in-themselves should be evaluative rather than factual and that the intellectual intuition must be admitted for the cognition of things-in-themselves.<sup>1</sup> Why? Mou answers as follows:

1. If the concept "things-in-themselves" were factual, it would be impossible to understand it. This is because if it were so, it would become a void concept that would belong to the other world. Since it is cognizable only by an infinite God with intellectual intuition, it plays a merely negative, logical function, like the logical concept of negating something. This further makes the distinction between phenomena and noumena empirical rather than transcendental. The negativity or the passivity of things-in-themselves could easily lead us into taking

---

1 Zhongsan Mou, 現象與物自身 (*Xianxiang yu Wuzhishen, Phenomena and Noumena*) (Taipei: *Xuesheng Shuju*, 1984). All translations from this book are mine.

all phenomena as noumena and viewing the distinction as a matter of degree rather than a matter of quality. Viewing them in this way, however, surely is not Kant's intention, so we cannot view the concept of things-in-themselves as factual.

2. If things-in-themselves are not clearly understood, both things-in-themselves and phenomena are not clearly understood. This is because phenomena and noumena are basically complementary. A loose understanding of noumena leads us to a loose understanding of phenomena. This, of course, does not mean that our theoretical or scientific cognition does not play the role of providing knowledge about the empirical world to the full extent. Mou surely admits the value of scientific knowledge. His intention is just that if we approach things only from the perspective of theoretical knowledge, we ironically cannot apprehend the nature of theoretical knowledge. This further makes impossible the meaning of such basic meta-judgments of theoretical knowledge as "it [an item of theoretical knowledge] concerns not things-in-themselves, but phenomena".<sup>2</sup> If we, Mou says, base our saying on factual sensitivity and understanding, what we can say is only that what we know is finite or dim, not that what we know is only about phenomena.<sup>3</sup> This shows that Mou here takes both a teleological and a holistic approach.

3. Mou claims that a clear understanding of things-in-themselves and further, a clear understanding of the transcendental division between phenomena and noumena, are possible only in that the concept of things-in-themselves is evaluative. The evaluative concept of things-in-themselves, Mou says, should be regarded as what is apprehended by intellectual intuition and therefore not as a postulate, as Kant thought of it. In Mou's understanding, the reason Kant viewed the concept of free will (a thing-in-itself for practical reason) as a theoretical postulate just for the sake of moral activity, rather than as something apprehended by intellectual intuition, was that Kant took the position of emphasizing theoretical reason over practical reason. Although Kant in general talked about the priority of practical reason over theoretical reason, he was inclined to theoretical reason or scientific reason by first dealing with the problem of cognition and postulating things-in-themselves as unknowable in his first *Critique*, then viewing them as a postulate or a hypotheses rather than as knowledge in his second *Critique*.

---

2 Mou, 17.

3 Mou, 12.

4. Mou and Kant would agree with each other on many points, especially in the realm of practical reason. They are in agreement in saying that practical reason is prior to theoretical reason because the former realizes things while the latter just cognizes things. Both Mou and Kant also pay attention to the fact that what is known is different from what is conceived<sup>4</sup> and likewise would admit that practical knowledge through intellectual intuition is not theoretical knowledge through sensible intuition. Also both Kant and Mou acknowledge that the concept of things-in-themselves as a factual concept is a transcendent one that cannot be cognized positively in any way, but should be approached in a practical way (5:55–6) and dealt with as an evaluative concept. They also would agree with the belief that our freedom is shown by our action according to moral law, not according to sensual desire and inclinations. Why would Mou, despite these agreements between them, be unsatisfied with Kant's conception of things-in-themselves as a postulate? Why did he call Kant's postulate void immanence (虛的內在) and darkness (冥闇)?<sup>5</sup> In fact, when Mou said that for the dignity of the human being, we should admit that he has an intellectual intuition that apprehends (in fact, realizes) moral law, did he not view intellectual intuition as a practical postulate, just as Kant did?

5. The reason Mou was dissatisfied with Kant's conception of postulate is that Kant's moral metaphysics remains in the realm of metaphysics of morals (道德底形上學) without being a genuine moral metaphysics (道德的形上學).<sup>6</sup> According to Mou, while moral metaphysics centering on morals establishes a metaphysical system—that is, the system of explaining beings—metaphysics of morals is just a metaphysical exposition of morals that analyzes moral concepts as a theoretical approach to the realm of practice. Kant's concept of a postulate in his practical concern is “sufficient as long as it does not contain any contradictions” (5:4).<sup>7</sup> Kant's work in practical reason is nothing but an analysis to clarify that human morality is disclosed through the categorical imperative, that human freedom is recognized through the categorical imperative's moral laws, and that freedom is the foundation for the establishment of those moral laws. In Mou's understanding, Kant's position is nothing

---

4 Mou, 62–3.

5 Mou, 64–5.

6 Mou, 92–3.

7 Immanuel Kant, *Critique of Practical Reason*, tr. Lewis White Beck (New York: Macmillan, 1956).

but the result of taking a cognitive and theoretical attitude toward morals, even though he proves freedom through his practical concern. Mou, therefore, says it is void for Kant to divide reason into its theoretical use and its practical use and thereby call the practical use of reason practical reason.<sup>8</sup> According to Mou, the term “postulate” is greatly imbedded with a cognitive attitude.<sup>9</sup> This is why Mou says “we should disclose directly the substance of morals just through the consciousness of morals and do not have to think about the idea of freedom indirectly through the category of cause.”<sup>10</sup>

6. The essential difference between a postulate and intellectual intuition is that while intellectual intuition in Mou’s philosophy can show what is and is not moral law in the real world, a postulate in Kant’s philosophy cannot. This is because for Mou intellectual intuition cannot be wrong. Mou also says real freedom and real autonomy should not contradict moral laws.<sup>11</sup> On the other hand, Kant’s conscience, to be postulated for a moral life, cannot determine what is moral law. Surely Kant’s moral law and his postulates of practical reason themselves cannot be wrong. This immunity from wrongness of Kant’s practical reason is possible because it is just formal, being purely rational and separate from the heart/mind. This separation of practical reason from the heart/mind led Mou to say the following about conscience, considered as the command from practical reason:

Conscience conceived only as sensitive capacity (with regard to this practical reason or moral law) can just sense duty, but as long as it can’t determine the duty, it necessarily does not know the origin of duty and therefore can’t realize the duty about the origin of which we do not know.”<sup>12</sup>

According to Mou, Kant’s conscience itself cannot be wrong, but it is precarious in that it does not tell us what is moral law. This is the reason Kant’s postulate possesses subjective necessity while Mou’s intellectual intuition does not.<sup>13</sup> In Mou’s view, Kant should have thought that moral mind must be the foundation of moral law. To say that mind is the foundation of moral law is to say that mind can comprehend the essence of things-in-themselves. Can the mind, however, really be im-

---

8 Mou, 81.

9 Mou, 65.

10 Mou, 62.

11 Mou, 77–9.

12 Mou, 69.

13 Mou, 69.

immune to error? Did Mou not also admit, as Yang-ming did,<sup>14</sup> that innate knowledge sometimes does not manifest itself in the actual world? In this regard, Mou's position would not be very different from Kant's because he also admits that the immunity of our moral mind to error can be thought only in the ideal world.

#### 4. Some Critiques against Mou Zongsan

##### 4.1 Why Should the Concept of Things-in-themselves be Evaluative?

I agree with Mou in saying that the transcendental division between phenomena and noumena requires the concept of noumena to be an evaluative one. I, however, would think that Kant's position of making the concept of noumena both factual and evaluative is better in many respects than Mou's position of making it only evaluative. Although someone like Mou might be dissatisfied with Kant's dichotomy, Kant's scheme has the merit, in fact, of freeing science from ethics. This is the reason Kant admits the concept of things-in-themselves as a regulative concept, at least in the realm of theoretical reason. In other words, the concept of things-in-themselves as a factual concept shows that, at least for the sake of cognition, we cannot use the faculty of understanding beyond experience. Mou distinguishes intelligible knowledge of the noumena by intellectual intuition from empirical knowledge of phenomena by sensible intuition. Mou says that, regarding the relationship between subject and object, the former is an undivided cognition and the latter is a divided cognition.<sup>15</sup> If so, Mou should have named the former differently, rather than treating cognition by intellectual intuition as genuine cognition. In other words, as Kant points out when talking about the transcendental ideas as ideas of practical reason, it seems to be more adequate to use words like "vindication" instead of "explanation", "cognition", and "penetration."

---

14 Wang Yang-ming, *Instructions for Practical Living and other Neo-Confucian Writings*, tr. Wing-tsit Chan (New York: Columbia University Press, 1963), 98–9, 134, 216–7.

15 Mou, 61.



## 4.2 Why not a Postulate?

Mou understands Kant's moral philosophy as abstract, logical, and void on the basis of his understanding of Kant's conception of a postulate. Is Mou's understanding justifiable?<sup>16</sup> Is Mou's attempt to portray his position as "immanent transcendence", in contrast with Kant's philosophy, justifiable? The reason Kant views free will as a postulate is that its existence is hard to justify objectively. This is also true in Mou's case, however, because he also accepts the faculty of cognition as intellectual intuition without any proof. Mou may say that the intellectual intuition he emphasizes is self-evident and so is an absolute fact of consciousness. Yet it seems that such evidentness and absoluteness can be found in Kant's "facts of pure reason"—that is, Kant's "pure consciousness about moral law" (5:6, 31–2, 42, 47, 55, 91, 104). Unlike Mou's thought, we rather can find in Kant's conceptions of respect for the moral law or free will the Confucian intellectual intuitions, such as the four moral sprouts, innate knowledge, or Confucian moral will. This is because Confucian moral substance can eventually be revealed in practice. In other words, we can draw "the realization of things", Kant's characterization of practical reason, only from postulates. If Mou is on the side of immanent transcendence, Kant also would be.

As said before, the difference between the postulates and intellectual intuition lies in the existence of fallibility. We can, of course, say that Mencius' four moral sprouts or Yang-ming's innate knowledge cannot be wrong. This is so, however, in the ideal dimension, not in actuality. In the actual world, the so-called four moral sprouts or innate knowledge can be wrong. The ideal interpretation of Confucian intellectual intuition makes us blind to the naturalness of the four sprouts or innate knowledge. If we allow a more natural and realistic interpretation, then the division between mind and principle that Mou denies becomes inevitable. The difficulty of Mou's idealistic approach is that it neglects the naturalness of the four moral sprouts and innate knowledge. Kant, on the other hand, can be said to have both an idealistic and a realistic attitude (5:105):

Nevertheless, with respect to our own subject so far as it knows itself, on the one hand, as an intelligible being determined because of its freedom by the moral law, and, on the other, as acting according to this determination in the world of sense, it is obvious that all this is quite possible. Only the

---

16 Mou, 9, 48.

concept of freedom enables us to find the unconditioned for the conditioned and the intelligible for the sensuous without going outside ourselves.

### 5. Cheng Chung-ying's Critique of Mou Zongsan

Cheng believes that Mou's idealistic approach to Confucianism is biased if we look at it from a broad perspective. In other words, in Cheng's view, Mou does not properly disclose the Confucian sensibility because he pays too much attention to comparing it with Kant's philosophy. According to Cheng, Mou describes Confucianism in a framework that sees the human person (or a Confucian sage) as an abstractly isolated individual who exercises pure reason, whereby both knowledge and ethics are regarded as a completed product and such dimensions of science, ethics, and aesthetics are taken as separate. In Cheng's reading of Confucianism in its so-called onto-moral-hermeneutical broad context,<sup>17</sup> however, human beings communicate with others unendingly and develop their characters through this process of inter-communication in the social community. Here in the social community is where human knowledge and ethics are realized through this concrete activity and where humanity's intellectual, ethical, and aesthetical dimensions are recognized as elements that need to be integrated with each other. In this respect, the sacredness of the Confucian sage should be seen not as static and transcendental, but rather as more flexible and dynamic. Although Mou in Cheng's reading takes intellectual intuition or infinite mind as having no distinction between subject and object, he still takes a subjective, personal, and substantial position with regard to knowledge and action. Cheng believes that since the human being is one who changes through unending interchange with others in the community, the sacredness of the sage or the capacity of intellectual intuition should be explained on the basis of such an understanding of the human being. In this respect, the four moral sprouts are identical not with "goodness itself", but with "expansibility toward the goodness".<sup>18</sup> In other words, the four moral sprouts are moving toward the infallible, but at the same time, Cheng admits that it is very subtle and fragile; he

---

17 Chung-ying Cheng, *Contemporary Chinese Philosophy*, eds. Chung-ying Cheng and Nicholas Bunnin (Malden: Blackwell, 2002), 357.

18 Chung-ying Cheng, "Theoretical Links between Kant and Confucianism", *Journal of Chinese Philosophy* 33.1 (March 2006), 8–10.

acknowledges our susceptibility to wrongness in every aspect. This realistic interpretation is convincing in a Confucian context, because such a concept of an ultimate being as Heaven and Tao in Confucian tradition is often compared to the omniscient Christian God, described not just as an external and transcendent absolute being, but also as an empirical, internal transcendental being. Furthermore, ideal moral life in the Confucian context is thereby interpreted as an unending interchange between Heaven and mankind in a dialectical development.

As Cheng suggests, if the practical reason of Confucianism is understood as a dynamic processes of development, the four moral sprouts or innate knowledge does not have to contain metaphysical substantiality, nor does it have to take the form of intellectual intuition for the sake of securing the practical purpose of morality. This interpretation made by Cheng does not impair the superiority of “the subjectivity” and of “the morality” that Mou wanted to secure in his moral metaphysics.<sup>19</sup>

Despite the merit of saving Confucian sensibility, Cheng’s interpretation still does not make clear how the unending process of mutual communication can guarantee the ultimateness of things-in-themselves, a feature Mou and Kant try to secure in their conceptions of intellectual intuition or infinite mind. Is it not more honest to admit our limitation as human beings, as Kant did, rather than to introduce intellectual intuition or an unending process of inter-communication among beings in an attempt to surpass that limitation?

---

19 Mou, 21–40.

PART X

Chinese Perspectives on Self-Cultivation



## 48. The Kantian Good Will and the Confucian Sincere Will: The Centrality of *Cheng* (誠, “Sincerity”) in Chinese Thought

A.T. Nuyen

### 1. Introduction

To a Western observer, there is something quaint about the frequent reference to sincerity in the social and political debates in East Asia and parts of South East Asia. It seems that in this part of the world sincerity is still regarded as a great virtue, and the lack of it, or insincerity, a serious vice. For instance, in their public slanging matches, China and Taiwan often accuse each other of lacking in sincerity. In a blistering attack on the former president of Taiwan, Chen Shui-ban, Guo Zhenyuan, a researcher at the Chinese Institute of International Affairs, said: “He is totally insincere about improving relations.”<sup>1</sup> Indeed, no sooner had Chen Shui-ban got himself elected than China denounced him as lacking in sincerity.<sup>2</sup> Whenever reunification talks between the two Koreas break down, the charge of insincerity is typically raised. During the “spy plane” incident, profound anti-US feelings culminated in the headline “US lacks sincerity” in the *People’s Daily*.<sup>3</sup> It seems we are supposed to see captured in this headline the outrage of the Government and the people of China. In the 2001 election campaign in Singapore, the then Prime Minister, Goh Chok Tong, demanded an apology from the leader of an opposition party who had accused the government of inappropriately lending seventeen billion dollars to Indonesia, a charge the government took most seriously, given its emphasis on honesty and its strong stance against corruption, but emphasized that apology would be accepted only if it was made with sincerity.<sup>4</sup> An apology

---

1 *The Straits Times*, Singapore, October 5, 2000, 15.

2 *People’s Daily*, May 20, 2000 (<http://english.peopledaily.com.cn>).

3 *People’s Daily*, August 14, 2001 (<http://english.peopledaily.com.cn>).

4 *The Straits Times*, Singapore, October 30, 2001, 1.

was subsequently offered, but rejected on the grounds that it lacked sincerity. In the whole episode, the charge of insincerity sat alongside charges such as being a cheat and a liar. Clearly, to be insincere is at least as bad as being a cheat or a liar.

The examples above amply demonstrate how important it is to be sincere, and how serious it is to be insincere in East Asian and some South East Asian countries. The same thing does not seem to be true elsewhere in Asia, such as Indonesia, or Thailand, or The Philippines: it seems to be a peculiar feature of Chinese culture, or societies with a Chinese background. It is certainly not a feature of social and political debates in the West. Indeed, as mentioned earlier, it is rather quaint to a Western observer. As Lionel Thrilling has observed, while the value of sincerity “became a salient, perhaps definitive, characteristic of Western culture for some four hundred years” after the West came to know how to use the word, the “word itself has lost most of its former high dignity.”<sup>5</sup> For Thrilling, the contradictions and paradoxes inherent in the depiction of sincerity in literature are combined with the philosophical influence of Hegel and the psychoanalytical power of Freud to result in the decline of sincerity as a virtue in the West. Indeed, it was Hegel, “a mind of great authority”, who “proposes to us the dismaying thought that sincerity is undeserving of our respect.”<sup>6</sup> Nowadays, Thrilling points out, when we hear the word “sincerity”,

we are conscious of the anachronism which touches it with quaintness. If we speak it, we are likely to do so with either discomfort or irony. In its commonest employment it has sunk to the level of mere intensive, in which capacity it has the effect that negates its literal intention—“I sincerely believe” has less weight than “I believe”; in the subscription of the letter, “Yours sincerely” means virtually the opposite of “Yours”. To praise a work of literature by calling it sincere is now at best a way of saying that although it need be given no aesthetic or intellectual admiration, it was at least conceived in innocence of heart.<sup>7</sup>

Thrilling’s account explains a Western observer’s impression that contemporary social and political debates such as those mentioned above are somewhat quaint. The fact that sincerity is still thought to be a key virtue in Chinese thought—the quaintness—can be explained if it can be shown that sincerity has a central role in Confucianism, and

---

5 Lionel Thrilling, *Sincerity and Authenticity* (London: Oxford University Press, 1972), 12.

6 Thrilling, 46–7.

7 Thrilling, 6.

that contemporary Chinese societies still retain a strong connection to their Confucian origins. The latter issue is largely anthropological, and I will simply assume that the connection exists. What I will try to show is that in Confucianism, sincerity is, as Thrilling puts it, “an essential condition of virtue”.<sup>8</sup> I will argue specifically that Confucians (and the contemporary East Asians and many South East Asians) take sincerity to be a key virtue because they take it to be not just a quality of utterances, such as an apology, or a condolence, but primarily a quality of the will itself. As such, the Confucian sincere will is equivalent to the Kantian good will. This comparison explains the centrality of sincerity in Confucianism.

## 2. The Centrality of *Cheng* (誠) in Confucianism

Against the suggestion that sincerity plays a central role in Confucianism, it may be said that textual evidence suggests otherwise. Indeed, it is hardly discussed in the *Analects*. By contrast, the various virtues, including the five “constant” ones, are discussed in all key Confucian texts. On the whole, commentators have not paid much attention to sincerity, compared with the extensive discussions of *ren*, *yi*, and *li*. This neglect is regrettable, as the omission of an extensive discussion of sincerity in the *Analects* is not proof of its insignificance.<sup>9</sup> Indeed, at *Analects* 1.8, we do find Confucius urging us to “hold on to faithfulness and sincerity as first principles”.<sup>10</sup> In any case, it can be argued that the *Analects*, being oriented more toward practical issues and more didactic in nature, simply assumes, or sidesteps, the metaphysical and epistemological background discussed in the *Doctrine of the Mean* and the *Great Learning*. As it turns out, the notion of sincerity is extensively discussed in these latter works. Given the fact that discussions of the virtues in the *Analects* should be read against the background of Confucian moral metaphysics and epistemology and the fact that at the core of such metaphysics and epistemology is the notion of sincerity, the sug-

---

8 Thrilling, 3.

9 For an exception, see Yanming An, “Western ‘Sincerity’ and Confucian ‘Cheng’”, *Asian Philosophy* 14 (2004), 155–69. Despite some similarities, An’s account of sincerity is quite different from the account given here. For instance, An does not discuss it in terms of the sincere will.

10 Quotations from the *Analects* have been adapted from various sources.



gestion that sincerity does not play a central role in Confucianism is mistaken. In the *Doctrine of the Mean* sincerity is given a metaphysical significance, while the account in the *Great Learning* has epistemological overtones. In the former, sincerity is the central concept in the Confucian metaphysics of the *dao* and of human nature, while in the latter it plays the key role in the Confucian moral epistemology and moral pedagogy. To understand the role of sincerity, we have to turn to these works.

In the *Doctrine of the Mean*, one's nature is taken to be the *dao* itself, the Way. It would appear that in this work Confucius gives us the same advice that Polonius gives Laertes in *Hamlet*: "to thine own self be true." However, one's own self, or one's nature is taken to be continuous with, or determined by, the *dao*, the Way, and sincerity is the means of realizing that nature, the means of being true to the natural self. Thus, Confucius has gone beyond Shakespeare in speaking of the true self not just in terms of psychological and mental states, such as feelings of love and hate, and knowledge of what is good and what is bad, but also in terms of the metaphysical nature that determines thoughts and feelings. For Confucius, it is possible for a self that consists of certain thoughts and feelings to be different from the true self that is determined by the *dao*. Indeed, while the self in its natural state is already in harmony with the *dao*, due to the various distractions in the world, the self in the world has diverged from the true self. Sincerity is a matter of being true to the latter, not to the former, as Polonius probably thinks. It is a capacity that the *dao* has endowed in us all, although we need to exercise that capacity in order to be true to the natural self, or to be in harmony with the *dao*. Thus, in the *Doctrine of the Mean*, sincerity is said to be "the Way of Heaven" while learning "how to be sincere is the way of man".<sup>11</sup> If one succeeds in the learning process, as the superior man does, one is said to be "in harmony with the Way."

In the *Great Learning* the first mention of sincerity can be found halfway through the short text in the form of "sincerity of the will." Here the notion of sincerity is explicitly applied to the will. The sincere person is more than someone whose utterances and conduct are in line with his or her inner thoughts and feelings, or in line with his or her moral inclinations. The sincere person is someone who possesses the

---

11 Wing Tsit Chan, *A Source Book in Chinese Philosophy* (Princeton: Princeton University Press, 1963), 107. The spelling of Chinese terms adopted in this work is retained whenever it is cited.

will to return to the true, natural, self. This is clear from chapter 6 of the “Commentary”: “What is meant by ‘making the will sincere’ is allowing no self-deception.”<sup>12</sup> “Allowing no self-deception” can be taken to mean not deceiving oneself about what one really is—in other words, being true to one’s own nature. This meaning probably follows Zhuang Zi’s linkage of truth (*zhen*) to sincerity. In the *Zhuang Zi* (chapter 31), we find the declaration “*zhen* is *cheng*.” On the surface at least, this meaning is the same as the Polonian imperative, “To thine own self be true”, and is consistent with the claim found in the *Doctrine of the Mean*, if we take the self to be the natural self determined by the *dao*. The rest of the commentary confirms this reading. More importantly, we find here the claim that one’s own true self is the moral self. Thus, since it is one’s true nature to “hate a bad smell or love a beautiful color”,<sup>13</sup> or generally to hate what is bad and to love what is good, and since the sincere will brings oneself into harmony with one’s true nature, the sincere person hates what is bad and loves what is good (i. e., is a moral person).

Going beyond the *Doctrine of the Mean*, the *Great Learning* gives an account of how one cultivates the sincere will. Cultivating it is one step in the series of eight steps leading to the all-important manifestation of the moral, or virtuous, character. The series begins with “the investigation of things” in order to extend one’s knowledge. With the extension of knowledge, one can make one’s will sincere. The sincerity of the will leads in turn to the rectification of the mind, and this leads to the cultivation of the personal life, resulting in the regulation of the family, thus contributing to national order and finally world peace. The person who accomplishes all this is a moral or virtuous person, one having a “clear character”. Embedded in the series of eight steps, sincerity seems to lose its central significance. However, this view is quite unfounded. To begin with, the positioning of sincerity in this way results from Chu Hsi’s re-arrangement of the original text that had sincerity of the will as the first in the series (without “investigation of things” and “extending knowledge”, as later “restored” by Chu Hsi).<sup>14</sup> Chu Hsi’s re-arrangement was later criticized by Wang Yang-ming, according to whom “sincerity of the will, without which no true knowledge is pos-

---

12 Chan, 89.

13 Chan, 89.

14 Chan, 89.

sible, must come before the investigation of things.”<sup>15</sup> But even in Chu Hsi’s version, sincerity could still be said to have a central position. Thus, in the series of steps arranged by Chu Hsi, sincerity of the will mediates between what is internal to the person (i. e., his or her knowledge) and what is external (i. e., his or her relationship with others in the contexts of a personal life, of the family, and of the state).

What we find in both the *Doctrine of the Mean* and the *Great Learning* is an account of sincerity in terms of the sincere will that has to be cultivated in order to return to, or to be in harmony with, one’s true self, understood as the moral self. On this reading, sincerity is crucial to being moral. This is consistent with the early meaning of sincerity. Thus, according to Sim and Bretzke, in the *Book of Rites* the term *cheng* is used to refer to “a disposition of ‘oughtness’ within the human heart”.<sup>16</sup> This old meaning suggests that *cheng* has a foundational role to play in the structure of a person’s moral character. Many later Confucians have offered readings of the self and the *dao* consistent with the account of sincerity given thus far. For instance, for Lin Yutang, “*dao*” stands for “the moral law”,<sup>17</sup> and since the *dao* conditions the natural self, the latter is moral in its essence. If so, the sincere person who is true to his or her natural self is *ipso facto* a moral being. Something like this interpretation is also evident in Wang Yang-ming’s account of sincerity in his commentary, *Inquiry on the Great Learning*. Indeed, Wang confirms that the sincere person “really loves” what is good “as he loves beautiful colors” and “really hates” what is evil “as he hates bad odors”; thus such a person is natural and naturally moral in the same way as we naturally “hate a bad smell or love a beautiful color.”<sup>18</sup> When a person has learned naturally to love what is good and to hate what is evil “then his will will always be sincere.”<sup>19</sup> Thus, Wang too takes one’s true nature to be moral in character and takes sincerity as what brings one into harmony with the moral order. Many other Neo-confucians also endorse this view of sincerity. For instance, Chou Tun-I regards sincerity as the foundational moral substance that determines our moral choices, saying

---

15 Chan, 84–5.

16 Luke J. Sim and James T. Bretzke, “The Notion of Sincerity (*Ch’eng*) in the Confucian Classics”, *Journal of Chinese Philosophy* 21 (1994), 179–212, at 180.

17 Lin Yutang (ed.), *The Wisdom of Confucius* (New York: The Modern Library, 1938), *passim*.

18 Chan, 664.

19 Chan, 664.

in his *T'ung-shu* (*Penetrating the Book of Changes*): “Sincerity (*ch'eng*) is the foundation of the sage.”<sup>20</sup>

### 3. The Sincere Will and the Kantian Good Will

The account above clearly demonstrates that sincerity plays a central role in the Confucian process of becoming moral. The centrality of sincerity can be seen even more clearly when we take it to be primarily a quality of the will and when we compare the sincere will with the Kantian good will. As Kant says (4:392–3): “Nothing in the world—indeed nothing even beyond the world—can possibly be conceived which could be called good without qualification except a *good will*.”<sup>21</sup> We have seen that the sincere will is what brings the self in the world back into harmony with the *dao*. As such, it is a thing good without qualification insofar as the *dao* is good without qualification. Indeed, since only the *dao* is beyond this world, “nothing even beyond this world” could be good in the way the sincere will is good. We can now follow Kant in his account of the good will and show that the sincere will functions in exactly the same way.

Following immediately from the statement cited above, Kant says:

Intelligence, wit, judgment, and the other talents of the mind, however they may be named, or courage, resoluteness, and perseverance as qualities of temperament, are doubtless in many respects good and desirable. But they can become extremely bad and harmful if the will, which is to make use of these gifts of nature and which in its special constitution is called character, is not good.

As mentioned above, much has been written on the Confucian “talents of the mind” and “qualities of temperament”, such as humanity, righteousness, propriety, wisdom, and trustworthiness; all of these are “doubtless in many respects good and desirable”, to borrow Kant’s words. The question is whether they are good and desirable “without qualification”. Many commentators discuss them as if they are. Indeed, Cheng Chung-ying regards righteousness (*yi*) as “the fundamental prin-

---

20 Chan, 465.

21 *Foundations of the Metaphysics of Morals*, tr. Lewis White Beck (Indianapolis: Bobbs-Merrill (1959), 10.

ciple of morality”.<sup>22</sup> To be “fundamental”, *yi* has to be good without qualification. However, if the account of sincerity and the sincere will given above is correct, we have to give a Kantian reading of the Confucian virtues such as *ren*, *yi*, and *li* and reject readings such as Cheng’s. We have to say that Confucians as well as Neo-confucians think *yi* and all other virtues are only good if they are exercised by a person with a sincere will.

The Kantian reading has ample textual support. As can be seen in the passage from the *Groundwork*, quoted above, Kant’s good will “in its special constitution is called character.” As it turns out, this is exactly how the sincere will is understood in the *Great Learning*: as constituting a good character. Indeed, the *Great Learning* opens with the claim that the “Way of learning to be great ... consists in manifesting the clear character”.<sup>23</sup> The clear, or good, character is what we must cultivate first before trying to be benevolent (*ren*), or righteous (*yi*), or proprietary (*li*). As in Kant, the good character, constituted by the sincere will, or the good will for Kant, is the foundation that other virtues manifest themselves upon, or the source that other virtues draw their value from. This view is endorsed by many neo-Confucians. As already stated, Chou Tun-I regards sincerity as the foundational moral substance that determines our moral choices: “Sincerity (*ch’eng*) is the foundation of the sage.”<sup>24</sup> Without the foundation of sincerity, the virtues themselves are nothing. Thus, Chou Tun-I claims that without sincerity “the Five Constant Virtues and all activities will be wrong” and all the esteemed dispositions “will be depraved and obstructed.”<sup>25</sup> One of Chou’s pupils, Ch’eng Hao, takes sincerity to be what preserves the virtue of humanity (*ren*) and *ren*, in turn, to encompass the other four of the five constant virtues.<sup>26</sup> While it is true, as mentioned earlier, that there is hardly any discussion of the sincere will in the *Analects*, the Kantian view is clearly implied in *Analects* 17.11, where we find the Master complaining that there is a tendency to take a virtue in isolation as something good “without qualification”: “Surely, when one says ‘The Rites, the Rites,’ it is not enough merely to mean presents of jade and silk.” This can be

---

22 Cheng Chung-ying, “On *yi* as a Universal Principle of Specific Application in Confucian Morality”, *Philosophy East and West* 22 (1972), 269–80, 126.

23 Chan, 86.

24 Chan, 465.

25 Chan, 466.

26 Chan, 523.

read as saying that it is not a good thing to offer presents of jade and silk without sincerity.

Extending further the comparison with the Kantian good will, we can say that the sincere will is both free and autonomous. In the case of Kant, the good will does not render the virtues good by relating them to something external to the agent, to a source of goodness outside the agent. Rather, it grounds virtuous acts in the moral law rationally conceived by the agent. In the same way, to be sincere is not to be true to some external ideal, to some standard imposed from the outside. Rather, the sincere will grounds virtuous acts in the goodness of the agent's own nature, and this in turn is the product of the *dao*. Differently put, the *dao* that the sincere will leads toward is the agent's own natural end. As pointed out above, the meaning of *cheng* encompasses being true to one's own nature. Thus, the *Doctrine of the Mean* states: "It is due to our nature that enlightenment results from sincerity."<sup>27</sup> Since what the agent has to be true to is the agent's own natural self, the sincere will that wills a person to be true to one's self is free. Since the goodness that the sincere will aims at is internal to the agent, the sincere will is also autonomous and not heteronomous. To act sincerely in the Confucian sense is to act freely and autonomously in the Kantian sense.

The Kantian model is also useful in helping us understand the process of cultivating a sincere will. To say that the ultimate goal for the sincere will is internal to the agent, that to be sincere is to be in harmony with one's own nature, is not to say that it is an easy matter to be sincere, as easy as doing something "naturally". It is true that Confucians believed we are born good. However, it does not follow that to be good requires no effort. The Kantian model tells us that morality consists in the struggle to be good, in knowing what maxims are good to will, and in making an effort to act according to such good maxims. With this in mind, we can understand why the Confucian texts on the one hand refer to sincerity as returning to one's own natural way, and on the other speak of the importance of *learning* to be sincere. As pointed out above, the discussions of sincerity in the *Great Learning* are heavily slanted toward moral epistemology and moral pedagogy. Since being sincere is not something one can achieve without effort, the "superior man will always be watchful over himself when alone" lest he fall into the way of the inferior man for whom "there is no

---

27 Chan, 107.

limit to which he does not go in his evil deeds.”<sup>28</sup> This point is reinforced in the *Doctrine of the Mean*, where it is said that while sincerity itself is the “Way of Heaven”, learning “how to be sincere is the way of man.”<sup>29</sup> Only a sage is “naturally and easily in harmony with the Way.”<sup>30</sup> On the Kantian model, the sage can be understood as an intelligible being whose will is perfectly good. Indeed, there is no difference between the perfectly good will of a purely intelligible being and God’s “holy will”. In the same way, the will of a sage is perfectly sincere, and insofar as it puts the sage in total harmony with the Way of Heaven, it is also “holy”.

If we are born good, if the *dao* is natural in us, why does the need for learning arise? Why are we more or less inferior people struggling to be superior rather than already sages? The answer, according to Wang, is that “while the original substance of the mind [the *dao*] is originally correct, incorrectness enters when one’s thoughts and will are in operation.”<sup>31</sup> Thus, the first task toward being sincere is to “rectify [the mind] in connection with the operation of ... thoughts and will.” However, other Confucians take a more Kantian line, whereby one’s goodness is corrupted by external influences. In the view of Li Ao, a forerunner of Neo-confucianism, it is because a person “has been darkened for a long time” that, in order to “recover his original nature”, one must guard “against depravity [and] preserve(s) his sincerity”.<sup>32</sup> On the Kantian model, we can take the “depravity” to be caused by sensuous feelings, by worldly inclinations. To overcome depravity, one has to try to make the will sincere rather than to follow inclinations. Thus, Li Ao writes: “If one is to stop [evil] feelings with feelings, that is to magnify the [evil] feelings.”<sup>33</sup> Li Ao’s reading of the *Doctrine of the Mean* receives support from many Neo-confucians. For instance, Chang Tsai confirms the duality of the self, consisting both of sageliness and a part that is dominated by feelings and desires, the two being kept separate by a lack of sincerity and united when the will is sincere: “When the Way of Heaven [or principle] and the nature of man [or desires] function separately,

---

28 Chan, 89.

29 Chan, 107.

30 Chan, 107.

31 Chan, 664.

32 Chan, 456.

33 Chan, 456.

there cannot be sincerity ... Sincerity is the way according to which heaven can last for long and is unceasing.”<sup>34</sup>

As we have seen, the sincerity of the will enables one to “investigate the nature of things” and understand the distinction between good and evil. Chang Tsai refers to this kind of understanding as that of the “higher things”. By contrast, to understand what is good as a pleasing feeling or as something that satisfies a desire is to understand “lower things”. With sincerity, one can understand the higher things and recover the good human nature. On the other hand, one who simply follows feelings and desires without sincerity will not find enlightenment: “Those who understand the higher things return to the Principle of Nature ... while those who understand lower things follow human desires.”<sup>35</sup> When an ordinary person succeeds in cultivating the sincere will, in building a “clear character”, he or she may be said to be complete. Thus, as Chenyang Li has pointed out, chapter XXV of the *Doctrine of the Mean* “explicates ‘cheng’ in terms of its root component which is also pronounced ‘cheng’ but means ‘completion’: ‘Cheng is self-completion (*Cheng zhe, zi cheng ye* 誠者，自成也).”<sup>36</sup>

#### 4. Conclusion

If the account of sincerity above is correct, and insofar as East Asians and many South East Asians understand sincerity in this way, it is easy to see why the accusation of lack of sincerity is a serious accusation indeed. It is serious because it impugns not the public face of a person but his or her inner self. To be accused of insincerity is to be accused of moral degeneracy, or depravity for Li Ao, something that is clearly worse than cheating and lying. As for the public face of the self, without sincerity, other dispositions that would be publicly virtuous, including humanity, righteousness, and propriety, are worthless, like bank notes not backed by gold reserve or national product. The insincere person is bound to be false to others. Prospects for cross-strait China-Taiwan relations, or North-South reconciliation in the Korean peninsula, are not good when accusations of insincerity are being made.

---

34 Chan, 507-8.

35 Chan, 509.

36 Chenyang Li, *The Tao Encounters the West* (Albany: State University of New York Press, 1999), 41.



I have made several comparisons of the Confucian notion of the sincere will with the Kantian notion of the good will in order to draw out the moral significance of the former. One further comparison is worth making. As is well known, Kant looks upon the moral law as awe-inspiring and admirable. Since the perfectly good will is absolutely obedient to the moral law, it too is an object of awe and admiration. Indeed, as pointed out above, the perfectly good will is morally on par with the holy will. In the same way, even though the ultimate end for the sincere will is something natural in us, it is still something that fills the mind with awe and admiration. For that end is the Way of Heaven itself. And since the absolutely sincere will renders one in total harmony with the Way of Heaven, it too is a proper object of reverence. Thus, with the Kantian model in mind, it is easy to see why the Confucian texts speak of the absolutely sincere will in superlative terms. For instance, the *Doctrine of the Mean* declares that “absolute sincerity is ceaseless. Being ceaseless, it is lasting. Being lasting, it is evident. Being evident, it is infinite. Being infinite, it is extensive and deep. Being extensive and deep, it is high and brilliant.”<sup>37</sup> Just as the absolutely good will is morally on par with the holy will, absolute sincerity is “a counterpart of Heaven”.<sup>38</sup> Just as a person with an absolutely good will is worthy of membership of the kingdom of ends, a person of absolute sincerity joins the ranks of sages and “is heaven”.<sup>39</sup> Not surprisingly, in speaking of the sincere will, the actual language of the Confucian classics is almost identical to Kant’s own about the good will. At *Mencius* 7 A4, Mencius declares that there is no greater delight than to be conscious of sincerity upon self-examination. At *Analects* 16.8, we find Confucius saying the superior man stands in awe of three things: the ordinances of the Heaven, great men, and the words of the sages. Insofar as both the great man and the sage possess the sincere will, we can collapse the three into two and borrow some of Kant’s famous words to render the two passages from the *Mencius* and the *Analects* as saying that two things fill the mind of the superior man with awe and admiration, the starry heavens above and the sincere will within. Judging from the political debates mentioned above, they still fill the minds of contemporary Confucian Asians with awe and admiration.

---

37 Chan, 109.

38 Chan, 109.

39 Chan, 112.

## 49. Desire and the Project of Moral Cultivation: Kant and Xunzi on the Inclinations

*Scott R. Stroud*

### 1. Kant on Desire and Evil

This essay will argue that Kant and Xunzi conceptualize desire as prone to disorder, and that desire plays an important role in ritual and community activity. The consistent picture of what Kant's moral theory aims at is simple—a community of agents who use their external freedom of action and internal freedom of end-setting in a harmonious, sustainable fashion. This is the theme of Kant's *Metaphysics of Morals* (*MM*, 1797), where the division between systems of right (*Recht*) and virtue (*Tugend*) seems to extend the sort of system pictured in the *Groundwork of the Metaphysics of Morals* (*GMM*, 1785) as the Kingdom of Ends. Ideally, rational agents will act and will in ways that are consistent, both intrapersonally and interpersonally.<sup>1</sup> Kant makes it quite clear in his various works that such a system of moral agents is based on reason. The question becomes: What is the value of the inclinations or desires for Kant?

The initial reading one gets of the moral value of inclinations is not good. In the *Groundwork of the Metaphysics of Morals* (*GMM*, 1785), Kant portrays inclinations as a source of temptation toward “immoral” willing. If one can either be determined by the moral law or by inclinations, and only the former is truly free, then the latter quickly become seen as a barrier to freely (and morally) acting. In *GMM* Kant labels the inclinations as “a powerful counterweight to all the commands of duty” that can be summed up under the concept of “happiness” (4:405). At a later point, the inclinations and their specific objects are labeled as “conditional” and not as comparable to the intrinsic value rational agents have as ends in themselves. He notes that “the inclination themselves,

---

1 Scott R. Stroud, “Rhetoric and Moral Progress in Kant's Ethical Community”, *Philosophy and Rhetoric* 38.4 (2005), 328–54.

as sources of needs, are so far from having an absolute worth, so as to make one wish to have them, that it must instead be the universal wish of every rational being to be altogether free from them [*gänzlich davon frei zu sein*]” (4:428). Here the inclinations are being portrayed as having a negative value, and one may quickly rush to the judgment that Kant is condemning the inclinations as bad in themselves. This may not be the case, however, as the “*davon*” preposition in the previously noted phrase can be taken in terms of “free from” their control, as opposed to “free of” them (i. e., without inclinations). On a charitable reading, then, Kant is not claiming inclinations are inherently bad; instead, he is claiming (1) they are not an important part of moral worth (the Good Will) and (2) they (and their conditionally valuable objects) seem to confound one’s attempt to will out of respect for the moral law.

The major change from *GMM* to *Religion within the Boundaries of mere Reason* (*RBR*, 1793) is that in the former inclinations are described as being immediately a challenge to moral willing, whereas in the latter the threat to moral action comes with the inclinations growing in strength, demand, and in unsustainable ways. Such a reading of the immediate burden of the inclinations, as well as the developmental nature of their threat, is given in the *Critique of Practical Reason* (*CPrR*, 1787), where Kant points out (5:118):

For the inclinations change, grow, with the indulgence one allows them, and always leaves behind a still greater void than one had thought to fill. Hence they are always burdensome to a rational being, and though he cannot lay them aside, they wrest from him the wish to be rid of them.

In this passage, Kant does seem to advocate the “extirpation of the inclinations”, as this is what is connoted by the verb “*entledigen*”. Thus, the inclinations seem to be given a negative valence in themselves when evaluated from the standpoint of moral willing: they are so burdensome that one wants to be free of them. This is clearly changed in his *RBR* account, which provides a nuanced reading of inclination and the innate “predisposition” to evil in human nature.

*RBR* marks an important addition to Kant’s ethical thought, one that is crucial to seeing the extent of inclination’s role in moral cultivation. In this work he explicitly begins to use the terminology of “*Wille*” and “*Willkür*” to stand for, respectively, the will as practical reason and the power of choice. Agents are taken as using their power of choice (*Willkür*) to incorporate incentives into the maxims that govern their ac-

tions, thereby creating their disposition (*Gesinnung*). Kant points out (6:24) that the freedom of the power of choice in regard to specific actions/maxims cannot be determined by

any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom).

Kant later identifies the two incentives in general as the moral law and self-love (stemming from one's inclinations). For humans, evil relates to the subordination of the former by the latter in the maxim that guides an agent. From where does this evil stem, and how do inclinations fit into this account?

Kant continues throughout *RBR* to maintain that, considered simply as a natural creature, the human being is not necessarily evil. There are two sources of support for this claim, the "natural" elements in his notion of human "predispositions to good" and his overt claims that the inclinations (*Neigung*) are not inherently evil. Before Kant advances the intriguing claim that human nature has a "propensity" to evil (6:29), he starts by listing three "original predispositions to good in human nature" (6:26), and proceeds to analyze each as a possibility to act in morally worthy ways or morally unworthy ways. The notion of "predisposition" (*Anlagen*) is defined as "the constituent parts required for it as well as the forms of their combination that make for such a being" (6:28). In order for humans to have vices related to their animality, for instance, they must have that potentiality or possibility; in other words, they must have a predisposition to good in that physical form of (normally good) self-love. The operative assumption in both the first and the second predispositions noted (to animality and to humanity) is that they are normally good but can be corrupted to account for the vices Kant discusses.<sup>2</sup> Without misuse, these are sources of legitimate moral duties, and hence, are good aspects of our "natural" constitution. The third possibility (personality) seems immune to any type of misuse: respecting the moral law as a sufficient incentive is never a vice, but can instead be heeded or not heeded by the willing agent.

As for the inclinations, Kant later argues that they are quite good when considered in themselves. At the start of Part II, Kant argues

---

2 Henry E. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), 149.

(6:58): “*Considered in themselves* natural inclinations are *good*, i. e., not reprehensible, and to want to extirpate them would not only be futile but harmful and blameworthy as well.” At least four lines of argument can be given for this position. First, it seems the inclinations can be imagined as good because all parts of nature are to be conceived of as good, or at the very least, amoral. In Section I of *GMM*, Kant appeals to the (reflective) teleological principle that all organs of an organism are best suited to a purpose by nature’s decree (4:395), and similar judgments about how we are to look at “nature” and its objects appear in the *Critique of the Power of Judgment* (*CJ*, 1790). The naturalness of the inclinations does not automatically make them evil, since such a term does not seem to apply to nature (given that it is without free choice) and since it appears (to us) to design its creations with the highest level of functionality possible. Second, Kant often discusses the inclinations as (in sum) composing the concept of “happiness” (4:405). Since we have an “indirect duty” to foster our own happiness (4:399), the inclinations must not be *prima facie* objectionable.<sup>3</sup> The third argument is given in *RBR* at 6:35, where he argues that the human’s sensuous nature and the inclinations it entails cannot be the source of evil, since agents themselves do not choose to have their specific inclinations: their existence is not something chosen, but is instead an aspect of our physical existence. A fourth argument is given in the same section, with Kant maintaining that the inclinations hold no direct causal relation to our will. Instead, the power of choice in humans is used to determine if they will act upon those inclinations as incentive, or on the moral law as incentive.

Only after this analysis does Kant argue for a “propensity to evil” in the human being. This presents a problem simply because it seems that the natural side of humanity is “predisposed to good” and the rational side (personality) is definitely oriented toward the moral law as incentive. Given this as foundation, the challenge would be to argue that humans are evil in some deep-seated, innate way. Kant, however, argues in Part I of *RBR* that humans have an innate propensity to evil (6:29). The use of “propensity” (*Hang*) is confusing, given that Kant wants to distinguish this term from “predisposition” (*Anlage*). He begins with a definition of “propensity”, explaining that by it he means “the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*), insofar as this possibility is contingent for humanity in general” (6:29).

---

3 Victoria S. Wike, *Kant on Happiness in Ethics* (Albany: State University of New York Press, 1994).

As he explains in a footnote (added in the second edition of *RBR*), this is merely a “predisposition to desire an enjoyment which, when the subject has experienced it, arouses *inclination* to it” (6:29). Like the predispositions to good, the propensity to evil marks a trait of human existence that can, but is not necessitated to, lead toward “evil” action. The main difference is in the valence behind such an incentive. As Kant notes, the propensity to evil highlights the probability that a certain experience (alcohol for an alcoholic, for instance) will, once enjoyed, grow and enlarge in terms of being a habitual desire—an inclination is fed by this act. This inclination to a specific activity is not initially present in its full-fledged form, but exists in potential form in the human. Given the right experience to nurture such a propensity, the inclination in question begins to grow stronger in terms of desiring more of that given activity and in terms of the strength it takes to resist such an urge. Given the development of this propensity, agents will have greater or lesser capacity to adopt the moral law as primary incentive in their maxims. Kant calls this the possession of either “*the good or the evil heart*” (6:29).

I take this opportunity to note what to agents seems like a phenomenological given: the “fact” that their inclinations have some “weight” or “pull” when it comes to their acting on them. Of course, the Kantian position is that the inclinations are not responsible for evil; instead, it is an agent’s free choice to incorporate them into one’s maxim as subordinated to the incentive of the moral law that makes said agent “evil”. Kant discusses this at numerous places and seems to highlight this point so as to not fall prey to the objection he anticipated in the *Critique of Pure Reason* (*CPR*, 1781/1787) at A554/B582. There he brings up the example of the “malicious liar” and notes that no matter how pathological his condition seems to be, it must always be presupposed that he can freely choose not to lie. While this notion of imputation is fleshed out in more detail in other works, it can be seen in *RBR*: the inclinations do not necessarily determine one to subordinate the moral law to them qua incentives. Why would this be a concern? Simply because if the incentive were the sufficient cause of incorporation, the agent’s spontaneous freedom (*Willkür*) would be destroyed. Agents must have true freedom concerning what to incorporate into their maxim; thus Kant sandbags against inclinations being the cause of their incorporation. What this account misses, of course, is a point Kant exemplified in *GMM*: inclinations exert some sort of force to be incorporated in a (morally) non-desirable fashion. This should

seem obvious to human agents, since they have most likely experienced situations where a putatively moral action was not in line with inclination, and resisting said inclination was difficult. Depending on the actualization of the propensity to evil, an agent will have an easier or harder time resisting non-lawful inclinations and their forces, in comparison to incorporating the moral law as primary incentive (over inclination/self-love). The agent's choice (*Willkür*) is an activity that holds a certain strength (always greater than 0 % in terms of *force*) against resistance (i. e., against inclinations, variable in *force* and in *direction* with or against the dictates of duty). One can even postulate that the strength of the moral incentive to incorporation can vary per individual, depending on cultivation of helpful sentiments, moral feeling, etc. The force of the moral incentive, as well as that of the “natural” inclinations, appears always to be of some magnitude (never 0 %). Humans always have the chance of acting on the moral law as incentive, and hence being responsible as a moral agent.

This fact of imputability need not be contradicted by the fact that incentives hold some force that *inclines* an agent to incorporate them into his maxim. The first “grade” (*Stufe*) of this propensity to evil is labeled by Kant as “frailty”, whereby the force of the moral law as incorporated is not greater than the force of inclinations in regard to the instantiating of certain agent-recognized moral actions (6:29). This seems to be the classic problem of volitional incontinence, albeit given a Kantian cast with the maxim/incorporation scheme. The second grade is “impurity”, whereby an agent manages to do the right action, but without holding the moral law as the self-sufficient incentive (6:30). Other incentives, presumably those of inclination, are needed to motivate one to act in accord with the dictates of duty. The third grade of evil, “depravity”, occurs when agents actively reverse the order of incentives incorporated into their maxim—in other words, they *actively* subordinate the moral law to the inclinations (6:30). This last step seems to be evil “at its best”, with its development completed. The first and the second grades address problems in an agent actively trying to be moral.

All of these grades involve agents having the choice to incorporate the moral law as pure, primary incentive over inclination, but as also failing in some regard. This failure seems to relate to a *propensity* to evil in that it is indicative of inclination's strength toward incorporation. The agent cannot or should not extirpate the inclinations, but should seek “rather only [to] curb them, so that they will not wear each other out but will instead be harmonized into a whole called happiness”

(6:58). Inclinations are susceptible to non-sustainable or conflicting development, so an agent must attend to how they are being strengthened through use and satisfaction. A crack cocaine addiction is not sustainable (for long), nor is it conducive to the satisfaction of other inclinations (or happiness as a set of all harmonizing inclinations) or the ability rationally to will action/ends in the future. It would be a prime example of an inclination that holds the (harmful) ability to gain in strength once fed, and this general trait of a human and their inclinations is noted by Kant as the propensity to evil. He acknowledges in a footnote that the inclinations do hold some force inclining the agent to incorporate them, and that they make “more difficult the *execution* of the good maxims opposing them” (6:59). He also notes that “genuine evil consists in our *will* not to resist inclinations when they invite transgression” (6:59). This will must be assumed to admit of some estimation of varying strength, as the grades of evil and the development of lawless inclination involve varying degrees of moral strength in response to them.

While this “force” analysis of will and inclination cannot *determinately* explain certain configurations of incorporation (since Kant maintains at *RBR* 6:59n and elsewhere that the freedom of choice makes such causal explanations impossible), it can usefully describe the moral experience of willing in the face of inclinations. Such a reading preserves moral responsibility in the fact that each agent always has the ability to choose against inclinations (no matter how “developed” in strength) and in favor of the moral law as primary incentive, all the while making sense of what it means to say that such an agent has a “propensity” to evil (in the likely development of force behind inclinations to grow out of easy control). Inclinations are good, but humans have an innate propensity to develop strong and ultimately unsustainable patterns of inclination-driven behavior if attention is not paid to how these incentives are organized.

## 2. Xunzi on Desire and Moral Cultivation

For the classical Confucian philosopher, Xunzi (310–219 B. C. E.), moral cultivation was the central topic for philosophical reflection. Following the lead of Confucius, he set to work elaborating on how humans could make themselves better and more in line with the *dao*. Where he differs from Confucius, however, is in his elaboration of inner psychological elements in human moral activity. Xunzi’s focus



is on rectifying or correcting human nature. What is the resistant element in human nature? Many have been tempted to say it is desire, largely because Xunzi clearly indicates that “Human nature is evil.”<sup>4</sup> Some now question this interpretation, doubting that Xunzi used the idea of willful acknowledgement and rejection of some pre-established good.<sup>5</sup> What is the moral value, then, of the desires for an “enlightened” reading of Xunzi?

It is tempting to say the moral worth of the desires is neutral, since the “evil” reading would attribute too rich of a notion of moral good to human nature in the first place. Yet, one begins to see how the desires have an inherent capacity to get out of control if not actively ordered. For Xunzi, human nature (*xing*) is spontaneous, responsive to outer environments, and effortless.<sup>6</sup> It features desires as part of one’s immediate, responsive, effortless reaction to the environment. The other crucial part is the mind (*xin*), a controlling or orchestrating faculty within the human.

For Xunzi, the desires are not *prima facie evil*. They do not involve a recognition of the good and a rejection of it. They are merely natural, immediately spontaneous, and effortless reactions to objects of experience. How they become morally important is in their propensities to lead to two sorts of conflict.<sup>7</sup> The first type of conflict is *intrapersonal*. One sees in Xunzi a concern for desires and how they can easily get “out of control”. What characterizes “control” in this sense? The idea of sustainability is vital here. A person’s desires could be easy to fulfill in the short term, but difficult in the long term. Additionally, one’s desires could conflict—satisfying one could preclude satisfying another, important desire. Desires demand moral cultivation and attention partially because they hold an innate propensity to grow out of control. This is why Xunzi puts the mind (*xin*) in a position of controlling desires for their long term benefit and order.<sup>8</sup> The second type of conflict is *inter-*

---

4 X 23.1a. Homer H. Dubs, “Mencius and Sun-dz on Human Nature”, *Philosophy East and West* 6 (1956), 213–22.

5 Philip J. Ivanhoe, “Human Nature and Moral Understanding in the *Xunzi*”, in T. C. Kline and P. J. Ivanhoe (eds), *Virtue, Nature, and Moral Agency in the Xunzi* (Indianapolis: Hackett, 2000), 237–49.

6 X 22.1b. Citations will be to chapter and section number in John Knoblock’s translation, *Xunzi: A Translation and Study of the Complete Works*, vols.1–3 (Stanford: Stanford University Press, 1988–1994), preceded by X.

7 X 19.1a.

8 E.g., X 22.5a.

*personal*. Desires are wide-ranging among humans, and not all desires of all individuals can co-exist. Hence, it is a very real possibility that humans will suffer great interpersonal conflict as their projects conflict or their success becomes mutually exclusive. Xunzi is clearly convinced that desires are not interpersonally sustainable if left unregulated.<sup>9</sup> Desires are inherently self-directed, and a system of such self-directed agents is bound to suffer conflicts or disorder. “Greed”, “envy”, and “hatred” are all bound up in the love of profit (*li*) that fundamentally besets all humans.

Thus, human desires (and *xing* in general) can be said to be bad (*e*) not because of any given content (*viz.*, a willful rejection of the good), but instead because they are ultimately unsustainable and undesirable at some point in time. Mind (*xin*) must regulate and order them primarily because of this propensity to disorder. How consonant are the readings of desire given by Xunzi and Kant on human nature? What new ways of understanding these two thinkers are opened up by considering them from this angle?

### 3. Orientation, Desire, and Moral Cultivation

Many differences separate the rich corpus of Kant from the elaborate work of Xunzi. But, if one views comparative endeavors pragmatically, then one can ask: What sense can be *made* of these two thinkers, especially as put into conversation with each other?<sup>10</sup> I want to advance three theses that can be extracted from this reading of Xunzi and Kant on the role of desire in moral improvement.

1. *The challenge posed by the inclinations/desires to moral cultivation is one of disorder, especially in a social setting.* Both Kant and Xunzi find that the inclinations are not *evil* in an ultimate way; that status is left by both thinkers to some “higher” faculty of human life and its consequent use. In Kant, this would be the exercise of choice (*Willkür*), and in Xunzi it would be the use of mind (*xin*). Here I want to comment on the two aspects of one’s desires (here including the Kantian term, inclination) that are highlighted in the preceding discussion. First, the paramount concern seems to be *social* order. When left unaddressed, desires

---

<sup>9</sup> X 23.1a.

<sup>10</sup> Scott R. Stroud, “Pragmatism and the Methodology of Comparative Rhetoric”, *Rhetoric Society Quarterly* 39.4 (2009), 353–79.

are bad precisely insofar as they become conducive to more disorder. Xunzi worries about them doing this through pursuing a limited amount of communal objects; a similar story undergirds Kant's discussion in *MM* of property as related to humans on the shared surface of the earth. One's wants can quickly outstrip the source of objects demanded. This becomes especially the case when it is the wants of the many, not merely the isolated individual. Second, both Kant and Xunzi foreground a characteristic common to all desires that leads to problems: they are always *self*-focused. They either regard something that gives one pleasure (think of something pleasing to look at, etc.), or they involve others in relation to oneself (jealousy, envy, etc.). In the former case, the limitations on the supply of the objects desired will often be tested by the sheer amount of people desiring them, thereby causing disorder and strife. In the second case, such social desires are self-focused insofar as they compare one to others in possessions, power, respect, and so on. Kant places a high priority on stopping these interpersonal desires, and thereby proposes in *RBR* a social solution to a social problem of evil: the universal church, or ethical community (6:94). Xunzi worries that these sorts of desire will lead to one abandoning formed ways of relating to others out of love of profit (*li*), leading to strife and harm.

Desires seem so harmful to the project of moral cultivation precisely because of this mix of their (1) inherent self-focus and (2) propensity to grow in strength. My reading of Kant has hopefully highlighted how the latter feature fits into his morally scheme: one's power of choice is swayed (but not necessitated) by the power of one's inclinations. As inclinations become entrenched or habituated, one faces more of a challenge *not* to incorporate those as subordinate to the moral law in one's volitional activities. Strong inclinations make it easy to subordinate the moral law in one's maxims, and harder to subordinate the inclinations to the moral law as supreme motive. Xunzi is also concerned about this, as he continually worries that bad choices, bad models, and tempting environments in the formation of bad habits will sculpt the sort of *xin* that will not instantiate the *dao*. The morally preferred endpoint is an agent who is *oriented* toward the community, and one's relations to the others present. This is diametrically opposed to what I call a self-focused orientation, the sort of orientation that humans' natural endowments incline them toward. Moral improvement is what counteracts this tendency; it is the sort of project that can result for Xunzi in ritual principles (*li*) being upheld and instantiated, and for Kant in a community of agents

respecting each other's internal worth and outer freedom. Both thinkers' endpoints seem to be systems of orderly and consistent use of the freedom of individuals. Self-focused orientations deny this sort of outcome, and thus must be resisted in education and schemes of moral cultivation.

2. *Moral cultivation focuses on altering a subject's orientation toward self, world, and activity.* Both thinkers seem committed to what I call *orientational meliorism*. This can be described as the purposeful and intelligent improvement of general orientations toward the self, the world (including other people), and activity to enhance future experience. Kant falls into this way of conceptualizing moral cultivation. Moral willing, on both the *GMM* and *RBR* accounts, concerns the *way* one orients oneself to others and toward the objects of one's desires. Do we treat ourselves and our desires as having more value or worth than the desires (and projects) of others? Or do we follow something like the Formula of Humanity as an End in Itself (FHE) and value others as equal to ourselves? *RBR* simplifies this story: Does this agent prioritize the moral motive of FHE over the individualized content of her inclinations (self-love, in general)? If the former is the case, this agent is acting in a universalizable, moral fashion. If all *actually* acted like this, the kingdom of ends would be instantiated.

For Xunzi, the project is the same: How does one alter the mind (*xin*) of another person such that they are enabled to correctly order their desires and actions? How does one help others become more like the gentleman (*junzi*) or sage? Xunzi is not simply talking about rectifying emotions or habits. At a fundamental level, he seems to be attributing some power of autonomy to *xin*. First, he points out that teachers and sages used this faculty of mind to create rituals (*li*) in the first place. Second, Xunzi claims that *xin* allows one to limit one's desires and to reflect on the meaning of one's desires.<sup>11</sup> Thus, the power of mind (*xin*) can order, reorder, and limit desires in the quest for orderly systematicity. The sages simply had the cultivated *xin* to do this for their society and its future generations in setting up ritual principles.

Contra Janghee Lee's account, one can see the same sort of ordering power of choice operative in both Kant's *Willkür* and Xunzi's *xin*.<sup>12</sup> Kant does not limit autonomy to a realm separate from inclination

11 X 22.5 and X 3.13.

12 Janghee Lee, *Xunzi and Early Chinese Naturalism* (Albany: State University of New York Press, 2005).

and sensuous nature; instead, autonomy in his mature moral works comes in the *mastery* of one's sensuous nature through one's rational choices of ends and means.<sup>13</sup> Lee's account is limited precisely because it confines itself to the *GMM* reading of moral willing; *RBR* and other later works give a rich view of Kant trying to wrestle with a rational being mired in a world of sensuous forces. Autonomy, in his later works, must involve desire. It simply cannot be opposed to and abdicating of all natural forces. Kant seems to share Xunzi's urge to master human desires and to refine them into orderly systematicity. Kant noted this in the Collins lectures of 1785: "insofar as it [freedom] is not restrained under certain rules of conditioned employment, it is the most terrible thing there could ever be.... If freedom is not restricted by objective rules, the result is much savage disorder" (27:344). The key to preventing this disorder is the right sort of moral character in an agent, the fortitude and knowledge required to be motivated by the moral law, not solely by self-love. This is also the basic program of Xunzi in regard to the moral improvement of *xin*.

3. *Moral cultivation involves a combination of (a) individual initiative and (b) manipulation of external environments to change individual orientations.* How do you change individuals' orientations from self-focused ones to orientations that recognize and respect social aspects of life? For Xunzi, the latter involved fellow-feeling (*ren*), as well as an understanding of social distinctions (*li*). For Kant, equality was a larger focus, but relationality in general is still there. An individual's projects, desires, and happiness all relate to the projects, desires, and happiness of other people. The morally cultivated individual will take all of these facts of the moral situation into account in using internal (end-setting) and external (bodily) freedom in social situations.

One reading of Kant is to focus on individual initiative—the individual must simply follow the moral law. The richer reading, drawing on Kant's work in the 1790's, postulates that education, social interaction, art, natural environments, etc. all must be intelligently manipulated to make it more likely the individual will *choose* to be motivated by the moral law qua incentive. In Xunzi, the same dual approach is evident: at places he extols the power of *xin* to regulate one's desires, yet at other

---

13 Paul Guyer, *Kant and the Experience of Freedom* (New York: Cambridge University Press, 1993); Robert B. Louden, *Kant's Impure Ethics: From Rational Beings to Human Beings* (New York: Oxford University Press, 2000).

places he exclaims: “It is the environment that is critical!”<sup>14</sup> I believe that like Kant, Xunzi places the ultimate necessity of individual reaction to desires at the center of his system; but, also like Kant, he realistically recognized that social problems demand social ways of melioration. One’s living among others is a *social* phenomenon; thus Xunzi and Kant pay attention to the environment (including social environments). These environments involve and shape the individual’s mental capacities (viz., choice and mind), and one’s independent use of these capacities in future situations is thereby influenced by those past experiences.

I will conclude this study with one of the prime examples of environmental manipulation in both Kant and Xunzi—that of ritual. Let us start with Kant’s example of communion, a typical religious ritual in his life-world. Why would Kant think of this as an appropriate means of meliorating one’s orientation (one that is typically dominated by self-love)? My argument is that communion is a means to such an endpoint primarily because it is an *instantiation* of that endpoint in the present. If one is concerned about individuals being isolated, not simply physically but morally, the ethical community advocated in *RBR* posits the solution of instantiating *non-isolation*. In other words, rituals such as communion *put* an individual into a certain social situation with certain valenced meanings, these being instances of the sort of relationality that is desired in *future* situations. Look at Kant’s description of this ritual (6:199–200):

The oft-repeated solemn ritual of *renewal, continuation, and propagation of this church community* under the laws of *equality (communion)* ... has in it something great which expands people’s narrow, selfish and intolerant cast of mind, especially in religious matters, to the idea of a cosmopolitan *moral community*, and it is a good means of enlivening a community to the moral disposition of brotherly love which it represents.

Taking communion together, or simply worshipping together, *instantiates* a sort of community. This representation is thus useful in convincing one to desire *this* sort of community in other aspects of her life, as well as to make that the dominant orientation of the present (versus one of self-love). By instantiating the end desired, one goes a long way to building the resolve to instantiate it again in the future.

Xunzi’s idea of ritual (*li*) involves the same sort of reading. Ritual serves as an *experience* of relationality—social distinctions and relationships among individuals are foregrounded in a ritual that one takes

---

14 X 22.5 and X 23.8.

part in (but not the only part). Whereas self-focus derived from desires places one at the center of attention, ritual displaces one from being the pure center of value and effort. There is a way that one ought to behave in a given situation, and this transcends one's wishes and desires. One's role is simply that—a role—and one has the choice to follow one's desires (like Robber Zhi, say), or to observe societal rules of ritual that foreground relationships. Ritual (*li*) nourishes the desires of individuals, but not in the unrestricted way that one would pursue if the desires themselves were in command. It is an orderly nurturing.<sup>15</sup> Why is this order or form important? Because simply following our desires and nothing else would inevitably ignore some very real differences in the world. Thus, Xunzi admires ritual because it is “the highest expression of order and discrimination” and it enshrines a recognition of important distinctions between classes of humans.<sup>16</sup> It also represents the forms of life experiences (birth, death, etc.) in sensible form.<sup>17</sup> Much like Kant's use of ritual, Xunzi seems to value ritual primarily because it is an *experienced* instance of what we should aim for at all times: a recognition of distinction and all that follows from that aspect of relationality. Focusing solely on one's desires can obscure that recognition. Ritual cultivates one's *xin* to be able to participate fully in one's relational existence.

There is more to my story of these two thinkers on ritual.<sup>18</sup> The central point is that both of these thinkers employ *end-instantiation*: the creation of some effect in the future by instantiating that end in the present. Ritual becomes a way to manipulate the social environment such that individuals find themselves taking a certain physical and mental orientation *now*. This orientation involves a new way of adjusting desires to social roles and other people, and it plays an integral part in the project of moral cultivation.

---

15 X 19.1a, 19.1b.

16 X 15.4 and X 19.1c.

17 X 19.4.

18 See Scott R. Stroud, “Ritual and Performative Force in Kant's Ethical Community”, in Valerio Rohden and Ricardo Terra (eds.), *Recht und Frieden in der Philosophie Kants*, vol.4 (Berlin: Walter de Gruyter, 2008), 143–56; Scott R. Stroud, “Xunzi and the Role of Aesthetic Experience in Moral Cultivation”, in Roger T. Ames and Peter D. Hershock (eds.), *Educations and their Purposes: A Philosophical Dialogue among Cultures* (Honolulu: University of Hawaii Press, 2007), 420–38.

## 50. Kant and Daoism on Nothingness

*Mario Wenning*

### 1. Original Nothingness

If one wants to depict the difference between Western and Eastern philosophy, the concept of nothingness provides a promising point of departure. Western philosophy displays a forgetfulness of original nothingness while Eastern philosophy has engaged with nothingness in theoretical as well as practical ways during its long tradition.<sup>1</sup> Despite its totalizing and simplifying tendency, there is some truth in this caricature of a West-East divide centered on the respective oblivion and valuation of original nothingness.

By “original nothingness” I do not mean a lack of quantity or the result of the negation of an affirmative proposition. Rather the term is meant to be a placeholder for an active yet indeterminate principle that cannot be expressed in terms of being a substance but transcends, or, if one does not like the language of transcendence, underlies our determinate world. It includes experimental phenomena such as forgetfulness and creativity as well as meditative attempts of emptying out one’s everyday stream of consciousness.

In Western thought this kind of original nothingness has been seen with suspicion ever since Parmenides argued that it is impossible to speak of what is not. Since original nothingness is simply not, since it lacks any determination or, as modern language philosophy would tell us, since it is an empty concept<sup>2</sup> resulting from an unwarranted substan-

---

1 For such an attempt, cf. Ludger Luetkehaus’s *Nichts* (Frankfurt: Zweitausendeins, 2003). Gi-Ming Shien defends the opposite thesis by arguing that “the full development of metaphysics both in ancient Chinese Daoism and in Greek philosophy culminates in nothingness.” See his “Being and Nothingness in Greek and Ancient Chinese Philosophy”, *Philosophy East and West* 1.2 (Jul. 1951), 16–24.

2 Cf. Carnap’s famous critique of Heidegger’s lecture “What is Metaphysics?": “Überwindung der Metaphysik durch logische Analyse der Sprache”, in *Erkenntnis* II (1931), 219–41 as well as Ernst Tugendhat, “Das Sein und das



tiation of the grammatical particle “no”, we should better drop it from the list of philosophically interesting concepts. Rather than creating pseudo problems by misusing language, we should confine ourselves to the legitimate use of negation and not postulate the existence of something that, by definition, is a no-thing.

In East-Asian philosophical traditions, on the other hand, the concept of original nothingness has enjoyed prominence and the paradoxes it gave rise to have been interpreted as productive paradoxes.<sup>3</sup> The Buddhist conception of Nirvana as the ultimate freedom from craving and suffering, the Daoist emphasis on the originally name- and form-less *dao*, or the Kyoto school with its emphasis on absolute nothingness<sup>4</sup> are just some representative examples that readily come to mind.

Western metaphysics has always preferred to inquire into the nature of being instead of nothingness. In spite of Heidegger’s reminder that the question of being, the *Seinsfrage*, has been forgotten, in Western metaphysics from its inception with the pre-Socratics, the concept of being has enjoyed prominence over that of nothingness. Although the Christian theologians of the Middle Ages conceived of the possibility of a *creatio ex nihilo*, the Aristotelian unmoved mover, and thus a presumably existing being, was always considered to be the originator of such a creative act out of nothing. In his *De rerum originatione radicali*, Leibniz raises the ultimate question of ontocentric Western metaphysics “*quod aliquid potius existit quam nihil?*” This question, usually translated as “why is there something rather than nothing?”, inquires into the meaning of existence in general. It aims at a justification of the existence of being as such. The very form of the question suggests that nothingness does not exist. Since Nietzsche, considerations of emptiness have been more frequent. What unites these considerations is that they see nihilism as a danger that needs to be overcome.

These representative examples illustrates that nothingness has at best been seen with suspicion, while it has often been simply suppressed.

---

Nichts”, in *Durchblicke: Festschrift für Martin Heidegger zum 80. Geburtstag* (Frankfurt: Klostermann, 1970), 132–61.

- 3 The idea that there are productive paradoxes in Asian philosophy is too often dismissed or simply ignored by analytic interpreters. Cf. for example Chad Hansen who claims that “if you concentrate on *wu*, you will be quickly dazed by philosophical puzzlement.” *A Daoist Theory of Chinese Thought. A Philosophical Interpretation* (New York: Oxford University Press, 1992), 221.
- 4 James W. Heisig, *Philosophers of Nothingness: An Essay on the Kyoto School* (Honolulu: University of Hawaii Press, 2001).

Eastern philosophies, on the contrary, came to the conclusion that nothingness is at least as essential as being. The question how to relate to nothingness has been a major concern for philosophy as well as meditative praxis.

My concern in this paper is thus not to trace a line of influence—this is not supported by the sources—but rather to point to parallel sensitivities that are, to be sure, developed in a different philosophical register. The obvious differences are understandable not only due to cultural backgrounds, but also due to the two and a half thousand years separating Kant from the classical Asian traditions. Furthermore, classical Eastern thought was operating within a paradigm guided by the assumption of cosmic harmony, while Kant was deeply influenced by the achievements of modern natural sciences that start from individual cases and aim at universalizability.

If my comparison of Kant and Daoism with regard to original nothingness nevertheless emphasizes similarities rather than differences, it is because I hope to narrow the gap between two specific episodes of Western and Eastern metaphysics rather than opting for the fashionable assumption of unbridgeable incommensurability.<sup>5</sup>

What does the shared sensitivity between Kant and Daoism consist in? I want to argue that this sensibility can be summarized by the following three, interconnected theses: (1) it is in the nature of human reason to search for the unconditioned; (2) this unconditioned cannot be conceived of in dogmatic metaphysical terms as something existing, but is a hypothetical non-entity; and (3) the insight into the illusory nature of claims to knowledge concerning the unconditioned does not lead to epistemic despair but harbors ethical consequences.

I will provide support for these theses by first focusing on Kant's critique of dogmatic metaphysics and then, in a second step, comparing what Vaihinger calls Kant's "philosophy of the as if" to the conception of nothingness we find in Daoism. Finally, I will argue that in both Daoism and transcendental idealism we witness a parallel transition from a metaphysics of non-existence to an ethics of regulative principles.

---

5 The incommensurability assumption is developed for example by David L. Hall and Roger T. Ames in their coauthored books *Thinking through Confucius* (Albany: State University of New York Press, 1995), and *Anticipating China: Thinking through the Narratives of Chinese and Western Culture* (Albany: State University of New York Press, 1995).

While the nature of the transition from metaphysical questions to ethical consequences is much more apparent in Chinese philosophy—which is why many commentators just see Chinese metaphysics as an appendix or prelude to its main concerns in ethics—it is also present in the transition from theoretical to practical reason in Kant. Let us first go one step back and call to mind Kant’s original starting point.

## 2. Kant, Metaphysical Need, and the “As If”

Kant’s position within the ontocentric history of Western metaphysics is not just that of a revolutionary but also that of an outsider that could provide for a bridge between West and East. From his astonishing early “Attempt to Introduce the Concept of Negative Quantities into the Wisdom of the World” (1763), he revealed an uncommon interest in breaking with a metaphysics of being and taking nothingness in the form of negative quantities seriously.

At the central transition before the Transcendental Dialectic that was to become essential for Hegel’s discussion of nothingness in his *Logic*, Kant distinguishes four kinds of nothings ranging from empty concepts (*ens rationis*) through empty objects of concepts (*nihil negativum*) and empty objects without concepts (*nihil negativum*) to empty intuitions without object (*ens imaginarium*). Although he does not explicitly raise the possibility of original nothingness—nothingness beyond the concept-object ontology—he strives to develop a concept of illusory objects, not as mere negations or privations pointing to a lack, but as playing a positive role. His theory of necessary illusions (i. e., non-existing necessary projections and postulates) was groundbreaking, and I will return to it shortly when discussing metaphysical ideas.

It would of course be crazy to assume that Kant’s Copernican Revolution that replaces traditional metaphysics with the study of necessary illusion, a study we could call “nihilology”, was due to Asian influences. Kant did not study Chinese sources first or even second hand.<sup>6</sup> Accord-

---

<sup>6</sup> It has commonly been assumed that Kant simply accepted the stereotypes of his day. Although there was an increasing interest in Chinese philosophy from the days the first accounts by Jesuits arrived on European shores, there was also a deep suspicion against the superstition that was attributed in particular to Daoism. Voltaire and Leibniz turned to China with an interest in the secular and rationalist nature of Confucianism. Leibniz corresponded with his Jesuit friend

ing to J. J. Clarke, Kant was “venturing little beyond common stereotypes when he referred to ‘the monster system’ of Laozi who, according to Kant, taught that nothingness was the highest good and who advocated a kind of perpetual tranquility in which all distinctions are annihilated.”<sup>7</sup> Clarke convincingly argues that Kant did not reveal the slightest interest in what he considered mystic Asian thought.

However, the apparent lack of sympathy or genuine interest that these remarks reveal stand in sharp contrast to some of the features of Kant’s transcendental idealism and, in particular, the role of nothingness within its architectonic. One could say that Kant’s attempt to limit traditional metaphysics from gaining positive knowledge of ultimate reality stands half way between the traditions.

Kant famously prefaces the A version of the *Critique of Pure Reason* by claiming that human reason is burdened by metaphysical questions that it cannot dismiss, since they spring from human nature, but that it also cannot answer, because they transcend its capacities. Metaphysics as the science aiming at ultimate foundations is a necessary aspiration of human reason, while reason can only create the illusion of providing answers about such ultimate questions.

Human reason thus tragically searches for an unconditioned conditioning ground that it can never know of, at least if “know” is understood as a theoretical activity with the goal of warranted propositional truth claims. Kant refers to this ground as *Ding an sich* or *noumenon*;<sup>8</sup> it is nothing for us, because in itself, it remains elusive for beings equipped with forms of intuition and a set of categorial modes of structuring our perception of the world that makes it impossible to experience the

---

Joachim Vovet out of a deep interest and respect for Confucian philosophy. He believed that in China one could rediscover a lost knowledge about a universal language that could bridge the warring religious and political factions of Europe.

7 J. J. Clarke, *The Tao of the West. Western Transformations of Taoist Thought* (London and New York: Routledge, 2000), 41. Cf. also Luetkehaus, *Nichts*, 719.

8 For the purpose of this essay I treat Kant’s thing in itself and the noumena as coextensive concepts, an assumption that Stephen Palmquist and others have rejected. See Palmquist’s *Kant’s System of Perspectives: An architectonic interpretation of the Critical philosophy* (Lanham: University Press of America, 1993), ch. 6. My contention is that the “as if” hypothesis that I will develop is a third alternative to the problematic “two worlds” interpretation of Kant and the more recent deflationary interpretations that try to dismiss the transcendental metaphysical baggage as unnecessary to understand Kant’s epistemic and moral insights.

formless conditioning ground underlying our experience *qua* being formless. Not only do we not know about the nature of *noumena*; no relation, even that of causation, between the *noumenal* world and the phenomenal world can be formulated, since all such relations would be irreducibly linked to categories and principles, and these have no validity beyond the world of phenomena. Referring to *noumena* as the ultimate reality underlying our experience is already an ambiguous way of speaking, since it assumes ontological hierarchies and categories such as unity and existence, yet these are reserved to structure the phenomenal world. Therefore, in the strictest sense of the word, things in themselves are the unconditioned non-beings that we cannot say or know anything about. Their only known property is that they are unknown to us. That Kant nevertheless speaks of such elusive things, or of such a “thing”, since the ascription of quantities also belongs to the phenomenal world, is due to an ambiguity in the set-up of his project, an ambiguity that many commentators starting with his first interpreters have tried to dissolve with varying degrees of success, usually by denying things in themselves an independent ontological status.

I would like to make an alternative suggestion to preserve and perhaps justify the apparent ambiguity behind things-in-themselves-talk. That we speak about this reality even if in ambiguous or even self-contradictory terms is due to Kant’s often overlooked assumption that philosophy may engage in making hypothetical assumptions that do not need to be cashed out in terms of knowledge claims. Hans Vaihinger’s *Philosophy of “As-If”* was the first systematic attempt to point to this essential feature not only of Kant’s philosophy, but of human forms of cognition in general.<sup>9</sup> Kant’s most explicit statement of his “as-if” approach is to be found in his hypothesis of the Copernican Revolution: consider the world *as if* it were what appears to us; and consider the *noumenal* as if it were what underlies such appearance, even though we cannot know anything about it.

The same “as if” model that allows Kant to distinguish between an unknown *noumenal* and a known phenomenal world reappears in Kant’s treatment of metaphysical ideas: God, freedom, and the immortality of

---

9 Hans Vaihinger, *Die Philosophie des Als-Ob* (Berlin: Reuter and Reichard, 1911); tr. C. K. Ogden as *The Philosophy of “As-If”* (London: Routledge and Kegan Paul, 1924). Cf. also Eva Schaper’s original attempt to apply Vaihinger’s thesis to the thing-in-itself problematic: “The Kantian Thing-in-Itself as a Philosophical Fiction”, *The Philosophical Quarterly* 16.64 (July 1966), 233–43.

the soul. The transcendental dialectic dealing with these theoretically unwarranted metaphysical ideas is an attempt to warn us against the Scylla of assuming that we can know (or disprove the existence of) the unconditioned and the Charybdis that we could dismiss the very question about its nature as meaningless. Dogmatic metaphysics assumed that things that cannot be proven are valueless. Kant's philosophy of the "as if" allows us to admit features of reality that we do not have any knowledge about and cannot prove, but that still turn out to be valuable and even necessary when conducting our lives.

### 3. Daoism on Nothingness

There are striking parallels between Kant's account of *noumenal* reality as an existence whose only predicate is that it does not have any predicate, thus condemning it to unrecognizability, and the conception of nothingness in Daoism. Daoism is centered around the concept of *wu*. In its verbal usage *wu* means "not being there", "not being present" as well as "not having." As a noun it denotes "non-existence", "absence", or "nothingness." However, while in most Western languages, these words come with negative associations, this is not the case in Chinese.

In the tradition of Chinese Daoism the nameless *dao* is what gives rise to the *dao* that is structured by the complementary principles of yin and yang. The determinacy of a formed world arises through the generation of these principles, a distinction created by the "Great Ultimate" that ultimately derives from the Ultimate; Laozi and his followers characterize the latter as *wu*. The principles *yin* and *yang*, where *yin* is the passive, female force and *yang* the active, male one, are structurally analogous to Kant's two stems of cognition, the receptive sensibility and the active synthesizing understanding, respectively.

The word "*dao*", referring to the reality that gives rise to these differentiations, has itself often been translated as "nothingness", based on statements such as "it is not the *dao* that can be named."<sup>10</sup> Laozi presents *dao* not as a concrete object that could be seen or touched, but as the conditioning ground that underlies the world of touch, sight, smell and, most essentially in our context, cognizability. Similarly to Kant's transcendental idealism, Daoism also speaks of something it claims it cannot speak about. The discursive intellect and its medium language

---

<sup>10</sup> Laozi, *Daodejing*, ch. 1.

is incapable of grasping the *dao*, the way, because its categories are too static and do not resemble the flow necessary to travel the *dao*.<sup>11</sup>

In the *Daodejing*, the great philosophical poem, its author Laozi uses nothingness to explain the generation of the phenomenal world: “all phenomena are derived from existence, and existence from ‘Nothingness’.”<sup>12</sup> Nothingness is also introduced as that principle that provides for a sense of perfection: “what things in existence are beneficial, it is because ‘Nothingness’ is applied to it.”<sup>13</sup> Nothingness is the underlying and perfecting principle that cannot be thought because all thought applies finite categories to the structurally infinite original nothingness.

Robert Neville has convincingly argued that nothingness should not be understood as a cosmological principle of origin but rather ontologically, or rather nihilologically, as the noumenal ground that underlies our experience and is also more fundamental than time and space.<sup>14</sup> However, reflections on original nothingness transcend ontology. The aim of the Daoists was never merely philosophical, but primarily practical in that nothingness was understood in relation to human thought and action.

Nothingness in the Daoist tradition encompasses experiential phenomena such as incomprehensibility, intentionlessness, and forgetting. It is essential in understanding creativity since creativity involves the generation of something new through an often incomprehensible process of intentionless creation. The *Daodejing* often speaks of the utility of emptiness. Only by following the way effortlessly, by doing nothing simply for something else, by not intervening (*wu wei*) is happiness achieved: “One does things noncoercively / And yet nothing goes undone.”<sup>15</sup> What this paradoxical formula suggests is that we should not even try to achieve happiness, for this would simply be another instrumental effortful action. Through becoming trained in emptying oneself of any instrumental concerns, one paradoxically gains the fullest state of being in the here and now.

Starting with Zhuangzi, nothingness is understood increasingly as an inner subjective principle. It explains intuition and spontaneity, as gen-

---

11 Arthur C. Danto, *Mysticism and Morality: Oriental Thought and Moral Philosophy* (New York: Columbia University Press, 1987), 101–20.

12 Laozi, *Daodejing*, ch. 40.

13 Laozi, *Daodejing*, ch. 20.

14 Robert Neville: “From nothing to being: The notion of creation in Chinese and Western thought”, *Philosophy East and West* 30.1 (Jan. 1980), 21–34.

15 Laozi, *Daodejing*, ch. 48.

erated from a source that remains below the threshold of consciousness. It is original non-being that invites us to be creative.<sup>16</sup> Creativity is the process of filling voids, and artists as well as children perhaps best exemplify the process of following the way freely. Arrows, toys, and brushes become instruments of self release that allow a person to follow a path of creative doing rather than to impose one's will through instrumental action on the world.

This emphasis on playfulness and creativity necessarily harbors a critical component. Zhuangzi suggested looking at our world not as finished but from the perspective of what Chad Hansen dubbed "skeptical rather than dogmatic transcendence".<sup>17</sup> Skeptical transcendence allows one to see the world as contingent manifestations emanating from an original nothingness. Daoism thus looks at the world "as if" it could be different, not because it knows what this different world would or should be, but because it knows that what exists is founded on nothingness.

#### 4. From the Metaphysics of Nothingness to the Ethics of the "As If"

We might think that the dilemma of conceiving of an entity or even the world as if it were nothing would lead to theoretical despair. In the case of Kant and Daoism, however, we witness the opposite. After the limitations they reached in their metaphysical investigations, due to the impossibility of claiming knowledge about a determinate conditioning ground, they consider a productive notion of nothingness in practical terms. Knowledge had to be limited to make room for creative thought and action. After the task of establishing a metaphysics of being failed, they attempt to reinstitute the hypothetical "as if" existence, or, in the case of Zhuangzi, non-existence of an unconditioned in the practical realm. The ethics of the as if thus replaces metaphysics as first philosophy.

In Kant's case the metaphysical ideas, though they cannot be proved or disproved, are justified for the purpose of action as regulative ideas. As regulative ideas they become the preconditions to engage in mean-

---

16 Rolf Elberfeld, "Kreativität und das Phänomen des 'Nichts'", in Günter Abel (ed.), *Kreativität* (Hamburg: Meiner, 2006), 520–33.

17 Elberfeld, 222.



ingful moral action. While we cannot conclude theoretically that there is freedom, as judging and acting beings we find ourselves in a world where we cannot but imagine ourselves as if we were free and thus responsible for our choices and actions. Justifying moral choices and actions is fundamentally distinct from explaining the structure of the world.

In the case of Daoism we live spontaneously as if the nothingness were guiding us although it is not doing so consciously. The sage is said to derive from the insight into nothingness a philosophy and, most importantly, the practice of non-interference. As is well known, Daoism's ethical thinking is not at all in line with Kantian moral philosophy. While the former stresses harmony between the principles of yin and yang, the categorical imperative defines morality as acting out of active duty alone rather than out of passive inclination. While Kant emphasizes the primacy of reason, Daoism supports the independent claims of creative human intuition and the body. While the spirit of Daoist ethical theory is negative in emphasizing emptiness and not taking action out of duty, deontology is indifferent to specific values and focuses on the procedure to generate universalizable maxims.

However, apart from these obvious differences, due to their parallel starting point of sharing a deep skepticism about the traditional metaphysics of claiming to know about the existence of transcendent entities, both Kant and Daoism share an emphasis on freedom. Freedom can only be imagined in a not fully determined world where original nothingness allows for creativity rather than a transcendent God or other being. While Daoist spontaneity is of course not the same as Kantian autonomy, both are premised on the assumption that we, rather than our societies or even our own history of values, determine what is right and wrong, as if we were free of these societies and values. Nothing in our nature or socialization is right *ipso facto* and immune from critique, and it might even be the case that the values of an entire community are wrong.<sup>18</sup> When we want to know what the right thing to do is, we should step back from our own private interest and universalize (i. e., raise the question about the significance of our actions or non-actions for the community).

---

18 Here I am following Chad Hansen's "Metaphysical and Moral Transcendence", in Bo Mou (ed.), *Two Roads to Wisdom? Chinese and Analytic Philosophical Traditions* (Peru: Open Court, 2001), 197–228.

One difference is important to point out. While acknowledging the merely regulative character of the principles guiding moral actions, Kant does not fully dispense with the traditional metaphysics of being. Moral action, for him, is generated by a metaphysical fact of reason (*Faktum der Vernunft*), whereas Daoism sees such transcendent facts as mere illusions that lead us astray rather than keeping us searching the way. Here a fruitful dialogue between Kantian and Daoist ethics is necessary.

To conclude, we have seen that both Kant and Daoism criticize a metaphysics that starts from the concept of being. Daoism assumes that what underlies our world is the nothingness of an unspeakable *dao*, while Kant postulates the hypothetical existence of things in themselves that are, for creatures like us, non-beings equal to x or zero in that they are structurally primordial in relation to our modes of structuring the phenomenal world.

Kant and Daoism both replace the primacy of a metaphysics of being with an ethics of the as if. Kant cultivates the assumption of the “as if” in practical terms to secure the possibility of moral action. Daoism develops an ethics of nothingness and spontaneity that sees the world as contingent and thus as a world that, if we would learn to follow the path effortlessly, could be a better place. Both Kant and Daoism thus converge in their insight into metaphysical indeterminacy and the resulting assumption of human freedom as the precondition of moral (non-)action.

# 51. Competing Conceptions of the Self in Kantian and Buddhist Moral Theories

*David Cummiskey*

## 1. Introduction

Kantians emphasize the separateness and distinctness of persons. Buddhists, on the other hand, emphasize the interconnectedness and commonality of all persons, and all forms of life. The starting point of Chinese Buddhist ethics is the universality of suffering and the truth of interconnectedness. For Kant, the autonomy of the will is the source of the dignity of humanity. For Buddhists, the core values are wisdom and compassion. They, too, are thoroughly interconnected and interdependent.

Since contemporary Kantians argue that the distinctness and separateness of persons is the key to a justification of deontological constraints, it is not clear how a Buddhist approach to ethics could incorporate such constraints in its normative theory. This essay focuses on the contrasting conceptions of the self in Buddhist and Kantian theories and argues that the Buddhist conception of the self is incompatible with contemporary Kantian deontology. This raises a puzzle about the Buddhist justification for the prohibition on harming and killing.

Our focus is on Chinese Tiantai (T'ien t'ai) Buddhism in particular, especially as represented by the *Lotus Sutra* and its doctrine of skilful means. In the end I conclude that a consequentialist virtue ethics, and an indirect consequentialist account of moral prohibitions (the precepts of Buddhism), provides the best interpretation of the Chinese Tiantai Buddhist ethical systems. Many consider the Tiantai school of Buddhism to be the first distinctly Chinese sect of Buddhism and the *Lotus Sutra* the highest teaching of Tiantai Buddhism. Although the interpretation offered here can be extended to all forms of Buddhism, the emphasis on the *Lotus Sutra* limits its textual basis to Chinese and East Asian Buddhism in particular.

It should be emphasized, however, that the methodological approach of this essay is not a scholarly study of classic Buddhist texts. Buddhist moral psychology and its conception of the self are our interest and focus. Instead of a textual study of the Buddhist cultural tradition, this essay aims at a *rational reconstruction* of Chinese Tiantai Buddhist moral philosophy. What is the best, the *philosophically most defensible*, version of the tradition? The goal is to explore the central concepts and develop an interpretation that fits with the core concepts but is also independently plausible.

## 2. The Problem

To begin, consider an interesting feature of all Buddhist ethics: the unconditional constraints on harming any sentient creature. The first precept of Buddhist ethics is a prohibition on harming and killing. A commonplace of contemporary Western moral theory is that constraints are paradoxical in that they prohibit infringement of the constraint even when infringing it would prevent more harm.<sup>1</sup> For example, one cannot kill to prevent killings. One cannot violate a constraint to prevent more violations of that very constraint. The puzzle or paradox is why it is impermissible to minimize wrongs or killings. It would seem that if my killing is wrong, then killings should not happen, and thus when all killing cannot be prevented, the less killing the better. Why is the focus on the agent's action (do not kill) over and above what the agent can prevent from happening (even more killings)?

A common answer to this question is to appeal to a Kantian conception of respect for persons to justify constraints. The idea roughly is that the wrong action is intrinsically wrong because it violates formal, rational constraints on justifiable action and/or that these agent-centered constraints reflect the special status of other persons as ends-in-themselves (and not means only). As Rawls first emphasized, Kantian constraints are based on *taking seriously the distinctness and separateness of persons*. Kantian ethics emphasizes the distinct dignity of persons, founded on autonomy of the will, thus setting humanity apart from the rest of nature. So here is our core question: without these types of Kantian founda-

---

1 For discussion of the paradox of deontology, see Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982).

tions, what is the justification for the Buddhist constrain on harming? Why is a Buddhist not allowed to harm to prevent more harm?

It may seem that a simple solution involves extending Kantian respect to all sentient creation. It might be argued that Buddhism simply treats all *sentient life as an end-in-itself*: the interests of all creatures count morally, so Buddhism simply extends the scope of Kantian deontological constraints. But this simple solution is clearly inadequate. To *avoid a consequentialist interpretation* of the Kantian idea that we should “treat persons as an end”, we need to focus on the *alleged distinctness of humanity* from the rest of creation, in particular on the *separateness of persons as autonomous agents*. A focus on promoting the interests of all simply does not justify deontological constraints, because the idea that “treating as an end” involves deontological constraints, rather than promoting interests, is itself based on the *alleged distinctness of humanity* from the rest of creation—in particular on the *separateness of persons as autonomous agents* each with distinct conceptions of the good.

This is a common point of contact between otherwise diverse contemporary Kantians. Robert Nozick, Thomas Nagel, Stephen Darwall, Christine Korsgaard, David Velleman, Thomas Scanlon, Thomas Hill, Paul Hurley, Frances Kamm, and many more Kantian moralists have argued that a Kantian conception of agency is a necessary condition for constraints. On the other hand, Peter Singer and many others consequentialists have argued that a more basic respect for all sentient interests, based on the commonality of suffering, grounds consequentialism.

Of course, consequentialists do support constraints on harming as essential means of promoting the overall good. In addition, consequentialists, like Buddhists, can also focus on character and motives. A consequentialist virtue ethicist claims that *actions are right when they reflect the motives, character-traits, or virtues that (indirectly) lead to the best possible consequences*. In the end, we will conclude that these alternative indirect consequentialist approaches provide a clear basis for a philosophically defensible form of Buddhist virtue ethics.

### 3. The Buddhist Conception of the Self (and No-Self)

Chinese Buddhist ethics focuses on character and moral psychology, and it thus has much in common with Aristotelian virtue ethics. The Buddhist conception of the self, however, is really at odds with an Aristotelian approach. The Buddhist conception of the self is thoroughly

anti-essentialist and really not at all Aristotelian. Indeed, the self in Buddhism is more akin to Hume's bundle theory of the self.

The self for the Buddhism is composed of five aggregates: (i) physical form, (ii) sense perception, (iii) emotions and feelings, (iv) cognition, and (v) consciousness; and these aggregates are embedded in the forces of karmic causality that cause rebirth itself. No core, or essential self, transcends and survives the flux of change and impermanence. The self, like everything else, exists only as a relational thing that is interconnected and thoroughly dependent on a web of relationships.

The doctrine of interdependent origination (or codependent arising) is the core metaphysical doctrine of Buddhism. It asserts that all of existence is essentially interrelated, interdependent, and interconnected. This is the heart of Buddhist philosophy: "*One who sees interdependent origination sees the Dharma, and one who sees the Dharma sees the Buddha.*"<sup>2</sup> The Buddhist doctrine of no-self applies the doctrine of interdependent origination to the self, and concludes that there is no essential enduring self. Although the causal integration and slow transformation of the elements of the self creates the illusion of an enduring self, the self is simply a momentary configuration of discreet, although causally codependent, changing elements. The self is essentially interrelated, interconnected, and interdependent on the rest of existence. There is really no unitary permanent self. Again, the self is constituted by physical form, sense perception, emotions and feelings, cognition, consciousness, and the forces of karmic causality that cause rebirth itself. It is the relation of these changing elements that constitutes what we call the self, and nothing more.<sup>3</sup>

The first small step on the long path to seeing the emptiness of the self involves recognizing human interdependence. Each person is dependent on and fundamentally connected to other people and to a particular community. Confucian ethics, with its focus on family and social relationships, is in this respect similar to Buddhist ethics. Buddhism and Confucianism involve a similar worldview, but Buddhism goes further. Buddhism extends relational thinking to all other persons and indeed, to

---

2 Quoted by John S. Strong in *The Experience of Buddhism*, third edition (Florence, KY: Wadsworth, 2008), 109. I have learned much from Strong.

3 On the doctrine of no-self, see Steven Collins, *Selfless Persons* (Cambridge: Cambridge University Press, 1982). In Western philosophy these issues are explored in David Hume, *A Treatise of Human Nature* (Oxford: Clarendon Press, 1896), and Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984).

all living things, and to the natural world. As a result, individualism, often considered the central insight of modern Western thought, is viewed as a fundamental delusion from the perspective of both Confucian and Buddhist thought. Confucian ethics focuses on our connection with our family and community. Buddhism is based on a deeper and more pervasive connection between all things. Indeed, the conception of the interdependence of all beings, and an ideal of boundless compassion for all beings, replaces the relational responsibilities of Confucian ethics.

Compassion toward all living creatures and equanimity of mind reflected in all of one's actions, reactions, and perceptions are the twin ideals of Chinese Mahayana Buddhism. Wisdom and compassion are the Buddha-essence or Buddha-nature. The development of wisdom and compassion is the essence of the Path and the Middle Way, but meditation is still the essential means whereby we develop ever greater wisdom and compassion. Of course, the final goal of Buddhism is release from suffering. We all want to be happy and avoid suffering. The insight of the Mahayana tradition is that the key to the release from suffering is developing both insight and boundless compassion. In particular, two insights are essential.<sup>4</sup>

The first is that cognition, emotion, and will are all interconnected. Let us start with the cognitive theory of the emotions. The emotion of fear usually has a clear cognitive content that includes the belief that something is dangerous or harmful. To take a simple example, fear of flying in an airplane includes beliefs about airplanes, flying, and danger. Fear of flying also involves the will in that it often includes a sense of losing control and thus a cognitive awareness of vulnerability. Some emotions may be more instinctual but most human emotions are laced with cognition. Without the underlying beliefs, it would not be the same emotion. Cognition also involves the will and emotion. If we do not take an interest in the objects of thought, we simply cannot concentrate and take in the information. Indeed, the more engaged and interesting something is, the more we can concentrate and the more we learn and remember. So emotions essentially include cognition and cognition presupposes affective engagement. It follows that one can change one's emotions, passions, and desires by changing one's beliefs and conception of reality. This is how insight can transform character.

---

4 See the Dalai Lama, *Ethics for the New Millennium* (New York: Riverhead Books, 1999), chs.6–8.

Second, emotion affects cognition in another important way. If we are angry or upset about something, we cannot concentrate and think clearly. Indeed, even one's capacity for perception is diminished by powerful emotions. The Dalai Lama calls the emotions that disrupt our mind the "afflictive emotions." Emotions like anger, hatred, greed, and lust generate powerful desires and unsettle our minds. Indeed, they distort our judgment, undermine our will, and ruin our sleep. Furthermore, when we act on these desires their "satisfaction" does not leave us satisfied at all. If a person is emotional in this way, they are never at peace. Of course, if one does not act on afflictive desires, the passions do not just go away; they remain and still corrode from within. Yet, if instead the person acts on the desire and expresses hatred, for example, there is a momentary release of aggression but the person is no better off. In addition, by expressing anger one will probably just harden an enemy, who may then retaliate in turn. It follows that we need to remove the emotion itself, and this involves a deeper transformation of the self. Restraint is but a first step in moral development that is aimed at undermining the cognitive basis of anger itself. The person without the afflictive emotion of anger does not suffer from its loss, escapes its bad karma, and is thus only a gainer.

Insight meditation aimed at anger would first help one internalize a deep awareness of the self-destructive nature of anger. It would also focus on the source of the anger and reveal its causes, its thorough interdependence, and ultimately its emptiness. Anger has a cognitive component and is thus focused on an object, but the object has no real essence and is itself caused by and dependent on a complex web of connections. As the essence of the object of anger dissolves in the face of reflection, so too does the anger itself. Anger is rooted in delusion and eliminated by insight. In this way, insight into interdependent origination transforms the cognition itself, and reveals that anger and hatred, and all afflictive emotions, are based on delusion and confusion.

Not surprisingly, the satisfaction of these desires founded on delusion leads only to more misery and suffering, both for oneself and for others. But insight must get *into* the anger itself; the mere abstract knowledge that anger is afflictive and self-destructive does not extinguish anger. Similarly, if I simply give someone the facts of airline safety, that does not eliminate the fear of flying. The person's orientation must be shifted both cognitively and emotionally, and this is a matter of fundamentally transforming the way one thinks and feels. Insight medita-



tion is the method whereby we transform our mind and heart and fundamentally shift our orientation and character.

The moral rules (or precepts) are a first step, but without insight and understanding, rules and restraint alone leave desires alive to fester within. It is thus necessary to also reflect on the causes and nature of anger and on the real nature of the object of one's animosity. It is only through greater understanding, and long practice, that established habits of thought and actions can be altered and reoriented.<sup>5</sup>

Cognition and emotion are interdependent, each relying on the other for the contours and content of the cognition and emotion. The embedded relations of the experience also shape and determine the experience. All of the inner aspects of the self are dependent on the outer aspects of the self and in the end the self itself is a relational construct dependent on the circumstance and interest that gives rise to the particular use of the self-reflexive concept: self. Beyond and above the relational aspects, there is no essential transcendent self at all.

So, too, there is no essential responsible agent that is untouched by the flux of experience and who is, in some deep sense, blameworthy or praiseworthy. But this is not taken to be an excuse for wrong action. Wrong actions, or unwholesome actions in classical Buddhism, harm both self and others and this is reason enough to forsake them. The doctrine of karma reflects a causal order where wrong actions rebound and harm the agent too because of the nature of the wrong. Wrong, unwholesome, actions tear at the very social relations that we depend upon. In addition, wrong actions flow from unruly passions, the afflictive emotions of anger and hatred, and delusions that need to be curbed, not encouraged.

The Buddhist conception of responsibility and justification for punishment are thus also at odds with the Kantian retributivist approach. For the Kantian, coercion is justified when it is a hindrance to a hindrance to permissible freedom. Punishment holds agents accountable for their wrongful actions that violate the permissible freedom of others, and as such punishment treats the punished as autonomous agents. A system of coercive law (Kantian justice) thus both maintains the distinct boundaries between autonomous agents, and also treats all persons with

---

5 Buddhist insight meditation thus has a mission that was lacking in the earlier yogic practices. Indeed, insight meditation (and mindfulness of interdependent origination) is also a way of life, in that one can adopt a meditative stance almost anywhere and anytime.

due respect as free and rational beings, as ends in themselves and not mere means. Without the Kantian autonomous agent, Buddhists must take a different approach to wrongdoing and punishment.

There is much in the Buddhist conception of no-self, cognition, and emotion worthy of sustained discussion. We have offered only a sketch of the main themes. The important point for now is simply that the relational conception of the self also transforms the related conceptions of agency, responsibility, and accountability.

#### 4. The Wrongfulness of Killing

With this basic outline of the Buddhist conception of the self in hand, we now return to our main line of argument and inquire as to the basis of the prohibition on killing in Buddhist ethics. The first precept of Buddhist ethics is the prohibition on harming and killing. From this one might simply assume that this prohibition includes the constraint on killing to prevent more killings. In the more general case of harm or injury, however, this is not the case. For example, I may cause a lesser harm, like amputating a limb, to prevent a greater harm, like death from gangrene. We also impose lesser harms on some to prevent greater harms to others. For the Kantian, harming one person to help another person calls for special justification because of the separateness of persons. The Buddhist rejects the essential separateness of persons and so an alternative justification is called for. For the Buddhist, the reasons for not harming others and for not harming oneself are essentially the same. So if we can harm ourselves to prevent greater harms to self, why is it that we cannot harm some to prevent greater harm to others?

As a focal point for the justification of the constraint on harming, we have the agent, the subject harmed, and the relationship between the agent and subject.<sup>6</sup> As should now be clear, the Kantian focus on the agent and autonomous willing (on a pure rational will governed by the categorical imperative) is not available for the Buddhist. Any focus on rational agency itself requires a more Kantian conception of the self. Similarly, if we focus on the subject harmed instead, we need to focus on the suffering caused and not a Kantian conception of the special status and dignity of the person harmed. Again, without a robust

---

6 The discussion that follows owes much to Shelly Kagan, *The Limits of Morality* (Oxford: Oxford University Press, 1989), see esp. 24–32.

Kantian conception of the subject, we do not have grounds for concern for the subject, other than preventing the harm itself. But if we focus on the harm itself, we see that killing to prevent more killings is likely to cause less harm to fewer victims, and so the constraint on harming to prevent greater harms cannot be justified by appealing to harm to the subject. So, that leaves the relationship between the agent and the victim as the only possible focal point of justification.

The relational approach, however, is equally problematic: if I kill to prevent killing, then there is less of whatever is bad about the relationship of killing. If the relationship of killing is bad, then more killings are worse. Since the Buddhist does not have recourse to the Kantian conception of rational willing grounded in the categorical imperative or respect for the dignity of autonomous agents, only the consequences of one's actions is left to determine their rightness.

Indeed, the Buddha's wisdom is sometime summed up in his Five Remembrances: the first three are that old age, illness, and death are unavoidable; the fourth is the impermanence of all things and all that I love; and the fifth is that *my choices and their consequences* are my only true belongings.<sup>7</sup> For Buddhism it seems that rightness must be based on consequences.

Of course, it is tempting here to appeal to the doctrine of double effect, distinguishing between intending harm and foreseeing unintended harm. The idea behind double effect is that I am especially responsible for the consequences that I intend and only secondarily responsible for the harms I foresee but do not intend. Foreseeable harms are permissible in the name of good intentions as long as they are not disproportional to the intended good. This distinction between intended and foreseeable harms, however, itself calls for justification; and the common justification offered appeals to Kantian rational agency and respect for Kantian subjects. As we have seen, these are not available to Buddhists. Furthermore, Buddhism is committed to an ideal of *universal compassion and universal responsibility*. We are supposed to develop a strong

---

7 *Upajjhatthana Sutra: Subjects for Contemplation in Anguttara-Nikaya (The Book of Gradual Sayings)*, vol.III (London: Routledge & Kegan Paul, 1973), 59–61 (Book 5, Sutra 57). The Pali cannon version: “I am the result of my own deeds, heir to deeds, deeds are matrix, deeds are kin, deeds are foundation; whatever deed I do, whether good or bad, I shall be heir to that.” Tich Nhat Han's version: “My actions are my only true belongings. I cannot escape the consequences of my actions. My actions are the ground upon which I stand.”

sense of compassion and responsibility for all sentient being. With this comes a primary intention to relieve suffering whenever possible and thus also to prevent wrongdoing wherever it happens. So the appeal to the doctrine of double effect is a nonstarter from a Buddhist perspective.

Although the Buddhist commitment to passivism is well known, it is important to see that passivism can easily be viewed as a means and not an end-in itself. We have already seen that anger and hatred aimed at others is self-destructive and typically misguided as well. In this sense, acting on these emotions is a mistake. Since violence is almost always rooted in afflictive emotions and delusions, it is almost always a mistake. Nonetheless, it is not the case that Buddhists reject all uses of physical coercion. Buddhist countries, and in earlier times monasteries, have police and armies. Buddhist mythology includes the Four Heavenly Kings, who guard the four corners of the world, protect the Buddha's followers from evil and preserve the teachings (the Dharma) of Buddhism.<sup>8</sup> Similarly, Buddhist rulers also used their armies to protect their country and the Dharma. Even Tibetan Buddhism led by the Dalai Lama, famous for its commitment to non-violence, owes its existence to an alliance with the Mongol warlords Genghis, Kublai, and Altan Khan. Mongol armies protected the particular Buddhist sect ruled by the Dalai Lama and raised it up to the dominant political position in Tibet. Indeed, the relationship was so close that the grandson of Altan Khan was actually the 4th Dalai Lama. Indeed, even the title "Dalai Lama" was itself bestowed by the Khans: *Dalai* means "ocean" in Mongolian and signified the Ocean of Wisdom manifest by the Dalai Lama. (This is also why Mongolian Buddhism is a branch of Tibetan Buddhism.) Lastly, the current 14<sup>th</sup> Dalai Lama fled Tibet *after* military resistance failed, though he now emphasizes that non-violence is the best means to achieve Tibetan cultural autonomy and reconciliation with China.<sup>9</sup>

---

8 For a full account of wrathful Buddhas and deities, dangerous protectors of the Dharma, and other symbols of the invincible power of compassion, see Rob Linrothe and Jeff Watt, *Demonic Divine* (New York: Rubin Museum of Art, 2004). For the Four Heavenly Kings see 107–8; a more easily accessible online reference is [http://en.wikipedia.org/wiki/Four\\_Heavenly\\_Kings](http://en.wikipedia.org/wiki/Four_Heavenly_Kings) (retrieved 2010-02-12).

9 For additional discussion, see my article, "The Law of Peoples and the Right to War: From Islamic Jihad to Buddhist Pacifism", forthcoming in Michael Boylan (ed.), *Morality and Global Justice* (New York: Westview Press).

In short, the Buddhist position is not that one can never use force and violence for self defense or national defense; rather it is that the use of force and violence is usually counter-productive and it is thus always the means of truly last resort. From a historical perspective, Buddhist passivism seems to be pragmatic and consequentialist, not absolutist. We will now see that the indirect, pragmatic, consequentialist status of moral rules is also supported by doctrinal considerations.

## 5. The Doctrine of Skilful Means

The prime text of Chinese Tiantai Buddhism is the *Lotus Sutra* and the core doctrine of the *Lotus Sutra* is the Doctrine of Skilful Means (also translated as Expedient Means).<sup>10</sup> The *Parable of the Burning House* perhaps best captures this core doctrine of the *Lotus Sutra*: to explain the many doctrines of Buddhism, and the evolution and advance of Mahayana Tiantai Buddhism, the *Lotus Sutra* uses a parable of a rich man whose house is on fire but whose children are inside playing and will not escape the burning house. As a skilful or expedient means, the rich man tells his children that there is a cart outside the house waiting for them, and he tells each that there is the type of cart outside that each child desires, a goat-cart or deer-cart or ox-cart. In joyous anticipation each child runs out of the burning house to seize the particular cart desired by each. Once outside the burning house, there are no carts. The rich man responds to his disappointed children, however, by providing each child the same fabulous ox-cart that is actually the best of all the carts, and this transcends the original, more limited desires of each of his children.<sup>11</sup>

This parable was used to explain the many distinct and often incompatible doctrines, practices, and sects of Buddhism that flowed into China from India. The parable tells us that the Buddha has provided different sects and doctrines to different people, but each with the goal of

---

10 *The Essential Lotus: Selections from the Lotus Sutra*, tr. Burton Watson (New York: Columbia University Press, 2002). Also see Michael Pye, *Skilful Means: A Concept of Mahayana Buddhism*, second edition, (New York: Routledge, 2003). On Tiantai Chinese Buddhism, see Dan Lusthaus, "Buddhist philosophy, Chinese", in E. Craig (ed.), *Routledge Encyclopedia of Philosophy* (London: Routledge, 1998); retrieved 2010-02-12 from <http://www.rep.routledge.com/article/G002SECT7>.

11 Watson, *Lotus Sutra*, 34–49.

helping people escape the burning house, the treadmill of afflictive desires and delusion that lead to suffering. Once out of the fire, the more advanced doctrines and True Path, or Vehicle of Enlightenment, can be revealed; these include the doctrine of skilful means, the idea that rules, precepts, and rituals are themselves skilful means to achieve the ever greater insight and compassion that leads to inner peace and happiness.

The elevation of the *Lotus Sutra* to the central text of Tiantai Buddhism is distinctly Chinese, first because it answers the problem of how to make sense of the many sects and competing doctrines of Buddhism introduced to China. Second, the conception of the Buddha as a benevolent father, who must use skilful means to lead his children to happiness, fits perfectly with the well-established Confucian ideal of the ruler (modeled on the benevolent father) who acts with skilful means for the benefit of his subjects and must be trusted and respected. Tiantai Buddhism is thus especially apt and well suited to the Chinese cultural tradition.

Returning to moral philosophy, we suggest that the doctrine of skilful means also applies to the ethical precepts of Buddhism. Clearly, many of the hundreds of moral precepts for Buddhist monks serve the end of developing inner discipline and restraint; these are prerequisites to virtue, rather than ends-in-themselves. Similarly, the point of insight meditation is transformative, as we have seen (although, as we also have seen, insight and virtue are inseparably interconnected). Perhaps the basic ethical precepts, like prohibitions on harming and lying, are also expedient means and secondary rules that help keep one on the right path, but are not absolute rules. Actually, the parable of the burning house already shows that justified deception is not a prohibited lie. The father misleads his children justifiably for their own good as the Buddha misleads his early followers to get them on the right path so that they can come to see the deeper truth.

Similarly, here is an interesting example of compassionate killing from the Buddhist canon that reflects the defeasible nature of all moral rules:

One night deities inform a Bodhisattva sea captain that one of his passengers is a robber intent on killing 500 merchants and stealing their goods. He realizes the robber will suffer many ages in hell for his deed, and that his only option is to kill the robber and take the bad karma on to himself. Accordingly “with great compassion and skilful means” he kills the robber.

But by willingly accepting the karmic punishment, the bad karma is canceled.<sup>12</sup>

There are several things to notice about this example of compassionate killing. First, killing is usually harming and rooted in aggression. In contrast, compassionate killing is rooted in Great Compassion both for the potential victims and for the potential victimizers. In principle, compassion can justify preemptive violence. Second, the sea captain is a Bodhisattva and has foreknowledge provided by deities. Although compassion can justify killing, only enlightened beings can have the virtue and wisdom to infringe such basic norms as the prohibition on killing. Third, as a corollary, it follows that ethical precepts are rules for the unenlightened. The conclusion is that the less enlightened should stick to simple moral rules but the more enlightened the being, the more judgment and compassion should guide one's actions in confronting difficult moral decisions. The right act will sometimes involve infringing rules for the greater good of all. Although transgressions still result in some bad karma, the good karma rooted in great compassion more than compensates for the infraction and harm caused.

The resulting position is strikingly similar to recent consequentialist moral theories.<sup>13</sup> For example, R. M. Hare argues that archangels, with perfect knowledge and perfect character, could follow direct consequentialist principles, but that simpler folks like us humans, "proles", need moral rules so as to do what is generally best. Of course, in fact we find ourselves, to varying degrees, in different contexts, and during different times of life, between the extremes of simple-minded proles and perfect archangels. So, too, for different people and contexts, moral rules can be more complex and refined; and in some cases of moral dilemmas, we should directly do what seems to be best overall. Peter Railton has also defended a compelling "sophisticated consequentialist" moral theory that incorporates the virtues of character into a broader indirect consequentialist ethical system. Similarly, Robert Adams has developed a character-based moral theory that he calls motive utilitarianism.

---

12 From Peter Harvey, *An Introduction to Buddhist Ethics* (Cambridge: Cambridge University Press, 2000), 135–6.

13 R. M. Hare, *Moral Thinking* (Oxford: Oxford University Press, 1981); Peter Railton, "Alienation, Consequentialism, and the Demands of Morality", *Philosophy and Public Affairs* 13.2 (Spring 1984), 134–71; Robert Adams, "Motive Utilitarianism", *Journal of Philosophy* 73 (1976), 467–81.

This approach leads us to the solution to the justification of constraints, precepts (and virtues) in Buddhist moral theories. Buddhist moral philosophy should be interpreted as a form of consequentialist virtue ethics. Roughly, *actions are right when they reflect the motives, character-traits, or virtues that lead to the best possible consequences*. The Buddhist conception of the self, with its account of the afflictive emotions and the interconnectedness of all things, shows that the best set of motives clearly will include a strong aversion and commitment to not killing or harming. This is the most plausible justification of Buddhist constraints. Unlike a Kantian conception of the distinctness and separateness of persons from each other and from the rest of nature, it is completely in synch with the Buddhist conception of the self. This consequentialist approach is also completely compatible with Chinese Tiantai Buddhism; indeed, it solves the problem of the basis of constraints for all forms of Buddhism.

We have seen that Buddhist precepts and virtues are not based on a Kantian conception of the self. We have also seen that indirect consequentialism fits nicely with Chinese Buddhism and it provides a clear basis for Buddhist ethical precepts and virtues. We can conclude that either Buddhist precepts are based on indirect consequentialist considerations or Buddhists must provide some other alternative to the Kantian and consequentialist conceptions of constraints. If one rejects our consequentialist solution, one must provide a new, alternative justification for Buddhist precepts.



## 52. What Is Personhood? Kant and Huayan Buddhism

*Ellen Y. Zhang*

### 1. Introduction

Personhood is a polyvalent concept that is ethically directed and philosophically embedded, yet notoriously difficult to define. In Kant's theoretic framework the person or personhood is intimately connected to his conception of self that entails the notions of identity, individuality, autonomy, and freewill. All of these notions seem irrelevant to Buddhism, since the Buddhist doctrine of *anatman* suggests a "voidness of personhood" that would disrupt the Kantian idea of self.

In this essay I shall discuss two different yet interrelated accounts of personhood with regard to self in the Kantian philosophy: a transcendent and a transcendental conception of personhood, both speaking of a rationally unified consciousness. Then I shall employ Huayan Buddhism as an example to explicate the Buddhist conception of personhood and discuss how the Huayan doctrine of *Dharmadhātu-pratītyasamutpāda* embraces a transcendental conception of personhood (in terms of an "enlightened" state of knowing and a recognition of one's individual relationship to the larger existence) but rejects a transcendent conception of personhood because Buddhism presents a distinct interpretation of the ideas of identity, individuality, unity, and totality.

### 2. Personhood: Transcendent and Transcendental

The word "person" in the Kantian framework suggests the notion of individuals with at least the following interconnected dimensions:

- (1) Identity: things that are individuals in terms of their own identity and integrity;
- (2) Autonomy: things that are individuals in terms of their own independence;

(3) Freewill: things that are individuals as initiators of actions and are capable of forming ideas and making choices.

Items (2) and (3) in this list are based on item (1), identity or self-identity. But what is the self here? A body-self? A mind-self? A spirit-self? An action-self? A person-self? Which “self” is the real self, a self that constitutes the locus of self-identity and subjectivity?

In the West, the question of personhood *qua* self-identity confronting Kant can be traced back to the philosophy of Descartes, whose well-known statement “I think, therefore I am” takes self-consciousness as the cornerstone of his philosophical argument. But the “I” here is not a total person with a physical body and action; rather it is the “mind”, the self-identical subject of the Cartesian *cogito*. The question of self-identity in Kant is presented partially as a critical response to Locke’s argument on personal identity. For Kant, personal identity is neither a mere memory of the past in terms of “sameness of consciousness”, as Locke had suggested, nor a pure form of the sameness of the mind in a Cartesian’s view.<sup>1</sup>

What then is “personhood” for Kant? The word “personhood” is an abstract, metaphysical term that entails the question of what gives us a coherence or unity of our occasional, particular experiences that is unique and recognizable. Furthermore, the concept of personhood plays a role in both the theoretical and moral sides of the Kantian system; thus it is closely related to evaluative judgment with regard to assigning duties, obligations, integrity, etc. As such, personhood is a propositional or cognitive issue that carries logical presuppositions in terms of human experiences. To answer the question of self, Kant introduces the “unity of apperception”, whereby he attempts to establish a constitutive connection between some forms of self-consciousness on the one hand, and experience of an objective world, on the other. The subject-predicate relation, for Kant, is crucial for defining the true nature of personal identity and moral duty embedded in a person.

---

1 In *An Essay Concerning Human Understanding* John Locke theorized that personal identity is dependent on “sameness of consciousness”, yet is not dependent on “sameness of substance” (ed. P. H. Nidditch [Oxford: Clarendon Press, 1975], 332–5). Kant is questioning the notion of “sameness of consciousness” in introspection. For more detailed discussion on this issue, see Karl Ameriks’ book, *Kant’s Theory of Mind: An Analysis of the Paralogisms of Pure Reason* (Oxford: Oxford University Press, 2000) and Patricia Kitcher’s essay “Kant on Self Consciousness”, *The Philosophical Review* 108.3 (1999), 345–86.

What is the connection of the subject to consciousness, or of the person to self? One of the crucial questions is whether there is an independently existing subject that both lies behind its experiences, apart from them, and is connected to the introspective intuitions of inner sense (the object of I), whereby the utterance of “I think” is generated. In order to answer this question, Kant makes a distinction between the transcendental self and empirical self, and sees the latter as contingent and non-abiding. He claims (B158): “I have no knowledge of myself as I am, but merely as I appear to myself.”<sup>2</sup> He also points out (A107): “Consciousness of self according to the determinations of our state in inner perception is merely empirical, and always changing. No fixed and abiding self can present itself in this flux of inner appearances.”

The idea that “the consciousness of self is thus very far from being a knowledge of the self” suggests that I know I am a self, even though I do not know exactly what the self is. The unattainability of the self, according to Kant, lies in the fact that we are conditioned by our experiences of the mere appearance of the self rather than the reality of the true self. Kant puts forth that introspective experience does not provide any evidence of the existence of an independently existing subject. But this does not necessarily lead to the conclusion that there is no such fixed and abiding independently existing self, even though the very conception of the “I” itself is conditioned by other outside experience. In other words, the only understanding I can have of myself is through my own self-consciousness. Self-consciousness, in turn, points towards an inner awareness of one’s own existence that gives rise to the conditions necessary for experience.

On the other hand, personhood as Kant conceives it has both a transcendent and transcendental status. Any experience involves a self-conscious being that Kant calls “apperception” as opposed to “perception”. Since Kant’s aim is to move beyond the traditional dichotomy between rationalism and empiricism, the question is how the world comes to be understood by me instead of how I can bring myself to understand the world. That is to say, the structure of our concepts shapes our consciousness and experience of objects, and consciousness becomes the subject of the experience of the empirical self. Personhood or personality, according to this line of thinking, is not just an outgrowth of reason, or a cognitive identification with the self, but also a “unity of ap-

---

2 Kant, *Critique of Pure Reason*, tr. N. Kemp Smith (London: Macmillan, 1953).

perception” (A366) that unifies all the sensory data. Meanwhile, human reasoning pushes us to think of our consciousness of self in transcendental terms in an idealistic way. This is why, says Kant, we must have a unity; the self is bounded by the temporal immanent as well as by the transcendent. In other words, it is a unity of the relationship between the noumenal and the phenomenal, rationally unified by consciousness. Kant thus rejects the traditional concept of self-consciousness and attempts to explain how reason determines the conditions whereby experience and knowledge operate.

The word “transcendent” usually refers to the primordial being that is underived from anything else, and ultimately to the being of all beings, the ground of all that is. For Kant, both transcendent and transcendental are connected to experience of self-knowledge rather than a pure reality; the former refers to something that exceeds or surpasses the limits of experience (i. e., supersensible or trans-sensible), the latter to the necessary preconditions of experience. One may be tempted to ask: what is the unified consciousness for Kant that presupposes the cognitive structuring of a rational being? Kant believes that what is first given to us is appearance and that when it is combined with consciousness, it is called perception. The transcendent concept of personhood constitutes “a high level conception” of what being a person involves, and the transcendental concept of personhood enables one to recognize the nature of those particular and empirical experiences. The subject-predicate relation, then, engenders the transcendent idea of a rationally unified consciousness and a cognitive structuring of self as the content of the concept of personhood. This Kantian idea has had a significant influence on the contemporary discourse on the philosophy of self, including the concept of self-consciousness in the theory of the “hermeneutical circle.”<sup>3</sup>

---

3 For example, in his essay “Language, Epistemology, and Mysticism”, Steven Katz, whose “contextualist position” is influenced by Kantianism, has argued that what is called “mystical experiences” in various religious traditions are different from one another because each cultural and social context that has generated religious experiences *via* self-consciousness and self-knowledge is different. See Katz’s *Mysticism and Philosophical Analysis* (Oxford: Oxford University Press, 1978), 22–74.

### 3. *Anatman*: A Person without Personhood

As is well known, one of the core teachings in Buddhism is no-self or *anatman* (無我, *wuwo*), the claim that the personhood of a person is nothing but a unity of five aggregates (五蘊, *skandhas*): form/body, sensation, perception, predispositions, and consciousness. It follows that what a causal relation or a causal process actually is in terms of the five aggregates points to a conceptual impossibility to defining personal identity. For the Buddhist, nothing permanently existent can be identified as “selfhood” or “personhood” over and above these five aggregates. The argument of no independently existing self is also connected to the Buddhist doctrine of *pratityasumtpada* (dependent/conditioning origination), the main base for the Buddhist doctrine of “emptiness” (空, *sunyata*). As dependent/conditioning origination indicates, self is ultimately “void” or “empty”. Yet emptiness here does not mean “non-existence” or “nihility”; rather it refers to the notion of a lack of autonomous self-nature (無自性, *nihsvabhava*). In other words, self does not have autonomous and abiding existence since nothing, including self and self-identity, has an autonomous self-nature. What is considered as self is the result of conditions (因, *pratyayas*), dependent upon which things come to be, and as such, one can only speak of self or self-consciousness relative to these things. Therefore, the idea of self only points to some kind of regularity of interconnectedness between events found in experience, but does not in itself spell out the logical distinctions or the degree of difference between self and other.

It is commonly maintained in the West that moral responsibility requires such concepts as personhood and individual autonomy, and this is particularly true for Kantian moral philosophy. Yet this understanding of personhood is not compatible with Buddhism, since the Buddhist idea of no-self makes the conception of atomic individuality impossible. Early Buddhist philosophers such as Nagarjuna (c.150–250) hold the view that what is conceived as self must be either identical with its components (i. e., the five *skandhas*), or else different from them. If self is identical with its components then any changes in the components is a change in the self; if the components are constantly changing—that is, they are different from moment to moment—so is the self that is identical with the components. But it is equally problematic if the self is completely different from its components, since no changes in its con-

stitutive components can then have any effect on the self.<sup>4</sup> As such, Nagarjuna contends that a conception of personhood is not the sort of thing we could possibly encounter in introspection, and that this is a good reason for denying there could be such a thing.

Huayan Buddhism, as a significant Mahayana tradition in China, accepts the view held by Nagarjuna. At the same time, Huayan Buddhism continues the discourse of *anatman* by framing the issue in terms of the whole-part distinction, an argument expounded by the early Abhidharma tradition. For example, a “chariot” analogy is used by Nagasena (c.150 BCE) to show that a chariot is nothing more than the wheel, the axle, the pole, etc., put into a certain order (i. e., the sum total of its components). At the same time, none of the individual components can itself be regarded as a chariot. What Nagasena tries to say here is that there are only parts but not the whole (chariot). Therefore, each part is dependent on the whole, and at the same time, contributes to the whole. Yet the question remains: Is there something about the chariot that is irreducible to its components, that is, something that can be called “charithood”? Or, is there any difference between the total sum of the chariot as a whole and the charithood of the chariot?<sup>5</sup>

In fact, Nagasena does not simply point out the whole is dependent on parts, but also contends that there is something about the whole that is irreducible to its parts. Meanwhile, there is a certain intrinsic relationship among of the components with their underlying cooperative unity

---

4 See Jay L. Garfield, *The Fundamental Wisdom of the Middle Way: Nanarjuna's Mulamadhyamakakarika. A Translation and Commentary* (New York: Oxford University Press, 1995). Note, however, that the concepts of “entity” and “essence” are quite confused in Nagarjuna’s framework, as it is not clear whether these concepts refer to continuities, or something individual (in the case of “entity”), or something collective (in the case of “essence”) that can be registered to explain experienced repetition of objects. Richard H. Robinson claims that Nagarjuna, in his attempts to negate the self-nature of all things, defines *svabhava* for its own purpose as a self-contradictory idea. As a result, Nagarjuna “may have been guilty of battling dragons of his own creation.” See Andrew P. Tuck, *Comparative Philosophy and the Philosophy of Scholarship* (Oxford: Oxford University Press, 1990), 59.

5 The part-whole argument refers, in fact, to two kinds of wholes: strong and weak. A strong whole is one whose parts cannot exist independently from the whole, whereas a weak whole’s parts can exist independently from the whole. For a detailed discussion on this issue, see Tao Jiang’s essay, “The Problematic of Whole-Part and the Horizon of the Enlightened in Huayan Buddhism”, *Journal of Chinese Philosophy*, 28.4 (December 2001), 457–75.

that constitutes the whole as a whole. One can talk about the sum total of the chariot in terms of their individual parts, the wheel, the axle, the pole, etc., but not the charithood of the chariot, due to the irreducibility of the chariot to its parts. In this sense, the chariot as a whole is neither identical to nor different from its components. This understanding of self can be seen in the Huayan tradition, but it attempts to avoid the problematic of the part-whole argument in Nagasena as well as the negative expression in Nagarjuna by postulating an all-embracing notion of “non-obstruction of everything.” This holistic method is called “totality” in Huayan Buddhism and described as *Dharmadhātu-pratītya-samutpada*, to be explained later.

Suppose personhood is associated with human experiences *via* those elements mentioned in the five aggregates such as sensation, perception, and consciousness, yet cannot be reduced to any of them; then how do we talk about the coherence or unity of personal identity? If we put Buddhism within the framework of a Kantian system, where personhood is a propositional issue that carries a logical presupposition in terms of human experiences, then should this logical presupposition be over and above each individual experience at the outset, or has it already been built into each individual experience? If we say that the logical presupposition is non-causal, it would be contradictory to the Buddhist view of dependent-origination. Buddhist philosophers might embrace a Kantian logical presupposition as an *upaya* or “skillful means” (方便, *fangbian*) in a philosophical discourse at the level of “the conventional/relative truth” (俗谛, *sudī*), but they would disregard it as ultimately empty. Moreover, according to Kant, *a priori* judgments are based upon reason alone, independently of all sensory/empirical experiences and pointing to strict universality, whereas *a posteriori* judgments are grounded upon sensory/empirical experience that is limited and contingent. Obviously, this kind of dualism between *a priori* and *a posteriori* would be problematic to Huayan Buddhism, as the latter aims toward an all-inclusive totality.

Similar to Buddhism, Kant also suggests a kind of “dependent-origination” when he speaks of the experience of consciousness and human knowledge. Kant’s self, in a sense, is not that different from Descartes’, when the latter argued that our knowledge of self is derivative of the presence of thought. Nevertheless, Kant’s treatment of the thinking subject as a transcendent concept aims at resolving the question of human experiences with regard to the relationship between experiences of consciousness (B75): “We need only say that there are two stems of

human knowledge, namely sensibility and understanding .... Through the former, objects are given to us; through the later, they are thought.”

Here Kant's distinction between sensibility and understanding points to the argument that such statements as “I think” always involve other concepts. Kant regards transcendent ideas of reason, like the transcendental categories of understanding, as innate in that a reasoning being is led by the reason one employs. But Kant is not sure how transcendence of consciousness can transcend itself. He seems to draw a line between experience itself and pre-experiential or post-experiential reflections, yet frames them in a dichotomic way: Reason does not apply directly to experience or to any object, but to understanding. As such, Kant holds the position that we cannot directly know the noumenal world as it is, since the noumenal world is beyond our knowledge. Nevertheless, Kant maintains that self-knowledge that depends on the transcendental categories to perceive of objects in the world still functions in a meaningful way. Like Kant, Huayan Buddhism postulates a similar skeptical view regarding self-knowledge and self-understanding through reasoning upon experience, because within the subjective-objective scheme, it would be logically impossible for the “I” that is the subject of experience to be at the same time its own object. Because of this very problem, Huayan Buddhism turns to another alternative for self-knowledge, that is, a pure experience without reasoning, or an unprojected consciousness, neither in the form of *a priori* nor in the form of *a posteriori*.

#### 4. Totality: World-As-Itself and *Dharmadhatu*

Before we discuss the Huayan view on experience, we need to explain the Huayan view on reality of the world—that is, the world of *Dharmadhatu*. What is *Dharmadhatu* (法界, *fajie*)? It is usually translated as a particular realm that is defined as a sphere or domain, where human activities and thoughts take place according to a frame of reference under specific conditions. *Dharmadhatu* also refers to reality from the perspective of the Buddha. In his treatise “The Mysterious Mirror of Huayan *Dharmadhatu*”, one of the centerpieces of Huayan Buddhism, Chinese master Du Shun (557–640), spoke of three domains of contemplations: the contemplation of True Emptiness (the manifestations of Dharma nature), the contemplation of non-obstruction between *li* and *shi*, and the contemplation of non-obstruction between *shi* and *shi*. Meanwhile,



Huayan holds to the principle of “simultaneous arising” (同時俱起, *tongshi juqi*), the view that all realms can exist simultaneously. Moreover, these realms do not obstruct or interfere with each other, yet they mutually penetrate one another in a harmonious unity as the principle of “simultaneous non-obstruction” (同時無礙, *tongshi wuai*) puts it.<sup>6</sup>

In the Huayan system, *shi* refers to actual, particular, temporal, and differentiated (i. e., what is phenomenal), whereas *li* refers to universal, eternal, and undifferentiated (i. e., what is noumenal). Yet the very idea of non-obstruction also points to a non-dual kind of duality between the phenomenal and the noumenal. Then how should we interpret this notion of the contemplation of non-obstruction between *li* and *shi*? Does the Huayan teaching here suggest a kind of transcendental self-knowledge that makes the experience of the non-obstruction possible? Obviously, for Buddhism, seeing or contemplating is neither a pure sense perception nor a mental thought but a meditative perception that points to a particular type of epistemology, wherein both sensibility and the thought process should be erased.

Huayan Buddhism attempts to resolve the tension between the phenomenal and the noumenal, comprehending the totality of the opposition between *li* and *shi* by employing the part-whole argument. It first emphasizes the mutual identity and inter-dependence of all phenomena (i. e., all things are interrelated and mutually dependent for their empirical and concrete existence). In other words, there is interconnectedness between part and part, that is, *shi* and *shi*. At the same time, there is interconnectedness between that part and whole, *shi* and *li*. As Master Du Shun says, the whole ocean (*li*) in its entirety is present in one wave (*shi*). “In each individual *shi*, *li* pervades in its totality.”<sup>7</sup> Ultimately there is no disjunction between appearance and reality, and as such to grasp the nature of *li* does not mean to uncover the noumenal reality obscured by the transcendently constituted veil of phenomenon, but to understand the interconnectedness between part and part, and part and whole. It is called “mutual interpenetration and identification” (相即相入, *xiangji xiangru*) in that one (*li*) is in all (*shi*) and all in one because of the principle of Dharmadhatu-Dependent Origination.

Meanwhile, *li* is understood as emptiness. Yet unlike the notion of *sunyata* expounded by Nagajuna’s Madhyamikan philosophy that is a

6 See Thomas Cleary, *Entry into the Inconceivable: An Introduction to Hua-yen Buddhism* (Honolulu: University of Hawaii Press, 1983), 84–92.

7 Cleary, 91.

logical concept inferred from the notion of dependent-origination, *li* in Huayan Buddhism is the object (content) of perceptive experience that transcends ordinary perception in a meditative (*samadhic*/pure) state. The concept of non-obstruction is also developed from the doctrine of emptiness, because only emptiness has no boundary by itself, thus there is no obstruction at all. In this sense *li* is not a transcendental judgment-form in a Kantian manner, but bears a transcendental implication in that it functions as a prerequisite or a built-in structure (i. e., the original relatedness) for conscious perception with its awareness of reality in its totality.

Note that Huayan Buddhism, unlike Nagajuna's negation, seems to have an essentialist bent in its conception of *li*. The idea of the noumenon is often used interchangeably with various other names, such as the Absolute, the Buddha Nature, and True Suchness that is supposed to transcend all phenomena. But at the same time, Huayan Buddhism makes a great effort to avoid falling into a dualistic trap. It insists that all phenomena are the manifestation of one immutable noumenon, and they are in perfect harmony with each other, like the different waves of the same ocean. This perspective, however, requires a perceptual transformation in that *shi* is no longer *shi*, perceived in the form of the conditioned, impure and mundane world of phenomena.

Such a perspective in Huayan Buddhism suggests an epistemological transformation—a new modality of knowing and experiencing, with a Kant-like “transcendental turn” but without the consciousness of self. Even though Huayan, like Kant, focuses on the notion of self-knowledge in the sense that the knower (i. e., the subjective) is primary rather than the known (i. e., the objective), it overturns the conventional subject-verb formula when regarding the subject as the knower and the object as the known. In other words, from the perspective of *Dharmadhatu* there is no such thing as “the subject objectifies” but a predicative relationship aiming at direct knowing, since the distinction between the subjective order of experiences and the objective order of events has been blurred in the process of experiencing totality.

Non-obstruction can generally be defined as the complete freedom from all bindings, especially the binding of the “subject-ego-logical” mode of thinking (i. e., the locus of self) that prevents one from seeing the infinite possibilities of interpenetrations among the realms, despite the variations in space and time. What would happen if we read the Huayan's idea of totality in a Kantian manner? Suppose there are fundamental preconditions for all experiences to take place; they are, as

viewed by Kant, transcendental in nature. This means the “unity of apperception” or the “synthetic operations of the mind” are transcendental in nature as well. It follows that self-consciousness is experientially aware of this “unity of apperception” since it is characterized by a mode of reflective activity *qua* consciousness of itself. Yet the consciousness of self or the ego-consciousness is exactly what Huayan attempts to dismantle, even though both traditions suggest in a different way the intuitive, the mystical, and the ultra-empirical. Therefore, the Huayan notion of totality is transcendental *sans* dualism (subject/object), for it opens up to a level of experience that is neither *a priori* (before) nor *a posteriori* (after). In other words, Huayan Buddhism speaks of experience in terms of the unity of experience and experience of unity, while retaining the transcendental realm with a non-discriminatory perspective.

### 5. The Practicality of Self as a Moral Agent

For Kant, the concept of personhood includes, in addition to rationally coherent and unified consciousness, the capacity for action. In his *Groundwork of the Metaphysics of Morals*, Kant sees the concept of virtue as a necessary practical idea of reason that can motivate a moral action, even though he emphasizes the necessity of grounding morality in *a priori* principles. Kant’s distinction between “what can be known” and “what can be thought” is more functional in formulating self as personal identity with a practical dimension in that it suggests a moral agent who acts in a rational way. Kant defines the task of philosophy as consisting in the examination of knowledge for the purpose of determining the *a priori* elements, and the determination of the rules for their legitimate application to the data of experience. That is to say, ultra-empirical reality is to be known only by practical reason, wherein one finds three things: consciousness, rationality, and the agent. With his onto-theological constitution of metaphysics, Kant attempts to utilize the transcendental to describe the relation of the self to the object-oriented world, as well as to our concrete relations with others in reality. As pointed out earlier, what Kant has in mind is a philosophy grounded exclusively on principles that are inherent in and revealed through the operations of rationality in the form of intuitive intelligence. The principle of morality, then, must be a moral law that is capable of guiding us to the right action in application to every possible set of circumstances (4:387–90). This kind of moral philosophy contrasts with empirical

moral philosophy grounded in *a posteriori* principles—that is, principles inferred through observation or experience. Here, we see a tension described by Kant between morality from above (i. e., principle-oriented) and morality from below (i. e., context-oriented), both, however, being action-based ethics.

The Buddhist moral philosophy, on the other hand, stipulates the ethical dimension of persons in terms of action that includes an intentional/volitional thought (意, *citta*) or consciousness-seed (種, *bija*), with this thought in turn producing an action (業, *karma*). In early Buddhism it remains controversial how an act has a moral autonomy that allows one to distinguish acts that lead to defilement and acts that lead to liberation. If no-self is maintained, it means one can characterize an act as either morally praiseworthy or blameworthy because the act itself carries with it its own workings that can be identified regardless of the agent or circumstance of its performance. Buddhist ethics in this sense can be categorized as deontological, and the framing of the ethical mechanics of *karma*, as operative through a continuum of mental states, characterizes the practical concern in Buddhism. Nevertheless, the intentional/volitional thought in Buddhism is not limited to reason or rationality; instead it also involves feelings and even unconsciousness. Kant, on the other hand, does not think moral feelings are reliable for furnishing a basis for universal law and thus insists morality should not be grounded in moral feelings.

As a sinicized form of Mahayana tradition, Huayan Buddhism speaks primarily of consciousness of the totality that is organic and holistic, transcending consciousness of the self. Yet to hold this view of totality does not lead to the negation of the role of an individual person and his/her specific moral obligations. Instead, Huayan maintains that only when one becomes all-knowing, experiencing the interconnectedness of everything, can one truly embrace the other and practice *karuna* (compassion). For Buddhism, morality is not a matter of following moral laws or fulfilling duties, but a volitional (i. e., [free-]willed) action that will generate causes and effects. As a result, the ethical self is also a non-egocentric *subjectum*: the word *I* means “here I am, responsive and responsible for everything”, since all sentient beings, as Buddhists see them, share a common experience and a common destiny. This is the Buddhist version of “duty” to others. Therefore, no-self in Buddhism does not necessarily lead to the statement that there is no moral agent who wills and acts; but there is no fixed selfhood or enduring self-identity. The denial of autonomy/self-existence as such should not leave one

with a sense of metaphysical or existential privation, a loss of some hoped-for independence and freedom, but instead it offers us a sense of liberation *via* the realization of the interconnectedness of all things, including human existence.

In his “Philosophy as Metapraxis” Thomas Kasulis has described a form of philosophical reflection that is devoted exclusively to problems surrounding the nature and efficacy of religious praxis. Kasulis calls this form of reflection “metapraxis” and argues that we need to distinguish it from other types of philosophical reflection, such as metaphysics, that problematize what stands behind or above religious praxis:

Religious praxis generally has either a participatory or transformative function. It participates in, to use Rudolf Otto’s term, the “numinous.” It is transformative in its improving the person or community in some spiritual way .... Metapractical reflection inquires into the purpose and efficacy of the practice in terms of these participatory and transformative functions. Something happens, or at least is supposed to happen, in and through religious praxis. Metapraxis analyzes and evaluates that happening.<sup>8</sup>

Buddhism is, as Kasulis puts it, a form of metapraxis that emphasizes the purpose and efficacy of the practice in terms of a participatory and transformative function: attaining an enlightened mind. In sum, the Buddhist conception of personhood is paradoxical in that it denies an ontological self, saying nothing exists as a permanent self outside the unity of the five aggregates, but accepts an ethical self (i. e., a self-other *persona*) that is embodied in a continuum of mental states with its continuous causal efficacy. This notion of moral action makes the connection between the Kantian and Buddhist conceptions of personhood possible.

## 6. Conclusion

Compared to Buddhism, Kant still remains very much a dualist, speaking of “inner” and “outer” experiences, of “subjectivity” and “objectivity”, and of the “world-as-it-appears” and the “world-as-it-is-itself”. As Robert Solomon has observed,

Kant answers the question of self-identity with a complicated theory of two selves, one of which is the subject of the phenomenal world, the other of which is an object in this world. The complication of this dual self theory is

---

<sup>8</sup> Thomas Kasulis, “Philosophy as Metapraxis”, in Frank Reynolds and David Tracy (eds.), *Discourse and Practice* (Albany: SUNY Press, 1992), 178.

both one of the outstanding contributions and one of the fatally weak doctrines of Kant's philosophy as a whole.<sup>9</sup>

Buddhism, by contrast, tries to avoid any dualistic structure of epistemology.

For both Kant and Buddhism, the conception of personhood points to both moral theory and practice. Despite its metaphysical difficulties (for it remains a question whether moral consciousness can be logically and objectively evaluated and described), the practical dimension of morality in both Kant and Buddhism has not lost all sense of the life, the activity, and the world of the ordinary person.

---

9 See Solomon, 23.



PART XI

Kant in Dialogue with  
Other Asian Traditions





## 53. Kant and the Buddha on Self-Knowledge

*Emer O'Hagan*

### 1. The Importance of Self-Knowledge

The exhortation “know thyself” brings to mind the value and significance of self-awareness in a thoughtfully lived life. Self-knowledge is a form of practical knowledge or awareness that connects one’s self-understanding as a particular person with a broader understanding of the human condition. There is moral value simply in the process of coming to know oneself, even partially and incompletely, because in doing so one better knows what it is like to be a human being, a moral agent, or to use Nagel’s phrase, one person among others equally real.

When cast as the narrowly epistemological issue of whether first-person access provides *a priori* knowledge of the content of what is thought, the question of self-knowledge loses its connection with the practical dimension that makes it morally valuable. Here I will not discuss the problems besetting epistemic accounts of the problem of self-knowledge and the attempts to save (or reject) first-person authority and agency.<sup>1</sup> Neither will I discuss Kant’s own epistemic problem: that there can be no truly accurate judgments about the moral self. Kant himself does not forsake the search for self-knowledge but insists that we have a duty to develop it. Instead I will focus on the problem as Kant discusses it in the later ethics and practical anthropology, where he explicitly addresses the problem of coming to have accurate perceptions and, by comparing his position with its Buddhist counterpart, I will show that Kant’s account of self-knowledge, while sophisticated, is unlikely to achieve its end.

Of course, given modern Western philosophy’s Cartesian heritage and its lessons, we should be wary of claims to know the self by simple introspective awareness. Indeed no plausible view of self-knowledge

---

1 For a clear discussion of these issues see Victoria McGeer, “Is ‘Self-Knowledge’ an Empirical Problem? Renegotiating the Space of Philosophical Explanation”, *The Journal of Philosophy* 93 (1996), 483–515.

would include the claim that we are infallibly aware of all of our own mental states. By “self-knowledge” I intend a fallible awareness of one’s present mental states that is in some important sense immediate or non-inferential and a capacity to speak of them with authority.<sup>2</sup>

Within Western moral philosophy (unlike the Buddhist tradition) the practical importance of self-knowledge has gone relatively undiscussed. Even among those philosophers thought to offer sophisticated accounts of the role of the emotions in moral life, such as Aristotle, there is little recognition of the diligence required to develop an accurate perception of oneself, or of the centrality of this project to the moral life. For example, in advising us on moral self-improvement, Aristotle assumes that we are accurate judges of our dispositions and goes on to use this judgment to correct character by applying the right sort of opposing pressure, much as one might apply opposing pressure in order to straighten a bent stick of wood.<sup>3</sup> Yet it is in Kant, a philosopher not renowned for his subtle analysis of affective states and their significance to moral character, that we find a clear recognition of both the difficulty and importance of self-knowledge.<sup>4</sup> For Kant, “know yourself” is the first command of all duties to oneself, and self-knowledge is to be understood as a judgment made concerning one’s moral perfection, in relation to duty (6:441). In his very brief discussion of self-knowledge, Kant writes that to know your heart is to know the quality of your will in relation to duty. This, he holds, will both dispel contempt for oneself as a person and will counteract egotistical self-esteem. In short, he claims it will have purifying effects on the will and, we may suppose, this is why he describes it as the beginning of all human wisdom.

Kant’s numerous important insights into moral psychology have parallels in Buddhist psychology. Here I want to highlight these parallels, indicating their larger importance in an account of moral development, and show that Buddhism offers a better and more useful conception of self-knowledge as a technique for moral improvement. I will argue that by framing moral self-knowledge entirely in terms of duty,

---

2 In this I follow Matthew Boyle. See his “Two Kinds of Self-Knowledge”, *Philosophy and Phenomenological Research* 77.1 (Jan. 2009), 133–64.

3 *Nicomachean Ethics*, tr. Terence Irwin (Indianapolis: Hackett Publishing Company, 1985), 1109b4–8.

4 For an interesting overview of Kant’s position on the difficulty of attaining self-knowledge, see Allen Wood’s *Kant’s Ethical Thought* (Cambridge, Ma.: Cambridge University Press, 1999), 193–202.

Kant distorts its value and makes it less likely that the egoistic tendencies he is trying to reign in will be uncovered. The Buddhist conception of self-knowledge, by contrast, being framed in terms of the awareness of suffering, offers a more comprehensive understanding of self-knowledge, not tied solely to detecting self-conceit.

## 2. Kantian Self-Knowledge

One of the most fascinating and underappreciated features of Kant's ethical theory is its recognition of the psychological complexity of human beings. Kant recognizes both that our moral judgments are continually threatened by corruption grounded in the human tendency to see ourselves in a good light, and that this tendency can have deleterious effects on moral character, and produce mistaken judgments about what duty requires. For example, when discussing the duty of beneficence Kant worries that the satisfaction a rich person may find in an act of beneficence may be a way of "reveling in moral feelings" and goes on to wonder, in the casuistical questions, whether "a rich man's help to the needy, on which he so readily prides himself as something meritorious" really should be considered beneficence at all (6:454). Kant's worry, that by reveling in moral feelings we can fail to see that we are simply doing what is morally required, is entirely apt; over time, were such feelings to go unchecked, they would corrupt our moral character. Kant thus recognizes that ignorance of one's moral constitution is a threat to its development, and his suspicion of our claims to know what we are like leads him to propose that we are morally obligated to hold our intentions up to on-going scrutiny by comparing them to the requirements of duty.

But the attainment of self-knowledge is complicated by the human tendency to disguise ourselves. Kant believes that when we become aware of being observed, we dissemble (7:121): "If a human being notices that someone is observing him and trying to study him, he will either appear embarrassed (self-conscious) and cannot show himself as he really is; or he dissembles, and does not want to be shown as he is." Accurate observations of ourselves are made difficult (perhaps impossible) because we do this even when we observe ourselves; we tend to discover in ourselves what we have ourselves unconsciously put on view (7:133). Because we are vulnerable to the opinion of others it is impossible to get a clear, undistorted look at human nature *per se*. Still Kant

believes we can, through practice, develop our cognitive capacities and refine our attention in order to gain a more objective view of things (7:131–2). For example, he describes the cognitive ability to ignore certain aspects of a representation such as a wart on another's face, a gap between the teeth, or a defect of speech (“even when the senses force it on a person”) as a very useful form of cognitive authority. The development of this power of abstraction requires practice in gaining command over one's representations, diminishing their capacity to divert one's attention and their seeming control over oneself. So even while Kant expresses skepticism about our capacity for accurate self-observation he still holds that one's attention can be trained and brought under one's control.

Furthermore, his own worries about the dangers of moral enthusiasm suggest that Kant was firmly convinced that some forms of life are more prone to self-deception and moral corruption than others. He claimed enough knowledge of human nature to know that the blindness of moral enthusiasm (as found in religious fanaticism) is a real danger; moral enthusiasts indulge in self-conceit and lose sight of their duties in the process. Moral fanatics, he claims (5:85), forget their obligations and focusing instead on merit, they are moved to action by pathological incentives and develop “a frivolous, high-flown, fantastic cast of mind, flattering themselves with a spontaneous goodness of heart that needs neither spur nor bridle and for which not even a command is necessary.” His stark appraisal of the inner life of the moral enthusiast itself indicates his conviction that certain ways of attending to one's experience are morally dangerous. Kant warns against “exhortation to actions as noble, sublime, and magnanimous” when they simply satisfy the demands of duty, noting that duty requires submission to the moral law but inclinations secretly work against it (5:85).

Self-knowledge is the first of all duties to the self, and of it Kant writes (6:441):

This command is “*know* (scrutinize, fathom) *yourself*,” not in terms of your natural perfection (your fitness or unfitness for all sorts of optional or even commanded ends) but rather 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.

He claims that moral self-knowledge will “dispel *fanatical* contempt for oneself” as a person, and that it “will also counteract that *egotistical* self-esteem which takes mere wishes—wishes that, however ardent, always

remain empty of deeds—for proof of a good heart.” It will make us less misanthropic and less prone to self-delusion. Two duties follow directly: impartial self-appraisal in comparison with the moral law, and the sincere acknowledgment of one’s inner moral worth (6:441–2). Kant claims it is only through the hellish descent into self-knowledge that the path to godliness can be found, suggesting that one’s evil will must be abandoned before a good will can be put in its place. Clearly, for Kant, the duty of self-knowledge is not a trivial matter.

Still, it is not clear what the duty consists in: what is it to know one’s heart and how does knowing one’s heart improve one’s will? The main problem for Kant’s account lies in the capacity for self-deception in self-judgment.<sup>5</sup> If the need for self-knowledge arises at least partly out of our bias in self-assessment, then how can self-knowledge serve as the solution to the problem? For this problem to be solved, some part of experience must count as trustworthy evidence; there must be some reliably accessible information about what my will is really like that I can come to detect. As Kant says very little on the topic of self-knowledge it is difficult to piece together a clear account that coheres with his other theoretical commitments.<sup>6</sup> I will not attempt such an account here; my focus is rather to compare Kant’s account of self-knowledge with Buddhist moral psychology.

Kant’s sophisticated moral psychology correctly recognizes the need for, as well as the difficulty of, attaining knowledge of the quality of one’s will. But when Kant appeals to duty to provide the theoretical frame for gathering self-knowledge, he lets his own reverence for duty obscure his otherwise impressive moral psychology. Kant’s worries about the corrupting powers of self-conceit are both exaggerated and at times unimaginative. He seems unwilling to grant that introspective awareness of one’s motives and other mental states, although fallible, can supply evidence of the content and quality of one’s will. By contrasting his view with Buddhist moral psychology, I shall offer a direction for showing that appropriately framed introspective awareness can

---

5 Nelson Potter describes self-deception as “the inner lie, by which we defeat morality in us, and thereby defeat ourselves”, noting that although a significant theme in Kant’s ethics, it is a “subterranean theme”. See his “Duties to Oneself in Kant’s Ethics”, in Mark Timmons (ed.), *Kant’s Metaphysics of Morals: Interpretive Essays* (Oxford: Oxford University Press, 2004), 386.

6 A thorough attempt is made by Owen Ware in “The Duty of Self-Knowledge”, *Philosophy and Phenomenological Research*, 79.3 (2009), 671–98.

provide the connection between self-knowledge and wisdom that Kant assumes.

Perhaps the closest Kant comes to identifying a trustworthy bit of experience is his claim that conscience is itself unerring (6:401):

For while I can indeed be mistaken at times in my objective judgment as to whether something is a duty or not, I cannot be mistaken in my subjective judgment as to whether I have submitted it to my practical reason (here in its role as judge) for such a judgment; for if I could be mistaken in that, I would have made no practical judgement at all, and in that case there would be neither truth nor error.

Conscience is unerring in the sense that it recognizes sincere moral reflection; moral judgment (whether perfect or imperfect) has been duly undertaken if one's reasons for action have been brought under the scrutiny of practical reason. The object of conscience is not the correctness of my practical judgment, but rather responsible moral agency: did I carefully consider what duty required of me? There can be no duty to have a conscience; having a conscience is a necessary condition for moral agency and thus subserves our duty to ourselves (as innate judges) as moral beings. Conscience, described as an internal court, operates for Kant as a "doubled self", with both prosecutor and the accused in the same person. The verdict of conscience concerns sincerity in the application of moral standards to one's actions, so one's conscience can be good even when one's actions are morally wrong. We don't have to get it right to have a good conscience, but we do have to try sincerely to get it right.

When Kant claims that conscience is unerring he means I cannot fail to be aware of the sincere submission of my subjective judgment to practical reason and hence it seems I can trust the conviction of my conscience (as an awareness of my diligence in reflecting on my duty). But is it really plausible to think conscience could serve as a ground for self-knowledge? Recall that the problem of moral fanaticism for Kant is not that fanatics are insincere; it is that they transgress the limits of reason, mistaking an image or felt sentiment of nobility for the requirements of duty. Furthermore, why, given Kant's hyperbolic concerns about the human tendency to put oneself in a good light, should the verdicts of conscience be trustworthy, while introspection of one's motives is not? Surely I can be deceived about whether or not I have carefully reflected and can find in my conscience self-congratulatory thoughts that I have put there for my own observation. My *certainty* that I have done well may indicate only that I have been pursuing my own interests in

the name of morality. The *Groundwork's* example of the shopkeeper shows that, in the absence of a casuistical process, one's sense of having reflected well is potentially delusional. Conscience alone, as inner certainty, will not solve the problem.

Allen Wood has suggested that we can make sense of Kant's willingness to trust the verdicts of conscience, along side of his views about human nature, only by understanding the claim that conscience cannot err in a very specific sense. According to Wood, by understanding all self-deceptive beliefs about conscientious judgment as formally speaking failures of conscience, we can grant that there can be no erring judgment of conscience, as in such cases there has been no judgment of conscience at all.<sup>7</sup> The claim that conscience cannot err makes sense if we conceive of conscience as a kind of success term for the judgment involved. But if we accept this interpretation then we lose the very feature of trustworthy experience needed to provide a ground for self-knowledge, for the issue has become terminological and so fails to identify any phenomenologically trustworthy vehicle for knowledge.

In order to avoid this sort of problem, it may seem tempting to try to understand self-knowledge as inferential knowledge, rather than as something more immediately available in one's experience. Because it is part of the point of Kant's duty of self-knowledge to keep us from self-identification with "mere wishes", we may instead turn to our external conduct to get clear assessments of what we are like. Just as accurate perceptions of our friends are had by dealing with them, perhaps self-knowledge requires reference to deeds, not merely opinions of what we are like?<sup>8</sup> Kant sometimes suggests self-scrutiny must be grounded not in the introspection of motives, but rather in scrutiny of one's deeds. By considering my life conduct, for example, I can evaluate how it has been improved by the adoption of certain principles, and so have cause to infer that my disposition has improved (6:68). This confidence in my improvement cannot be based "upon an immediate consciousness of the immutability of our disposition since we cannot see through to the latter but must at best infer it from the consequence that it has on the conduct of our life" (6:71). So on this view, judgments

---

7 Allen Wood, *Kantian Ethics* (Cambridge, Ma.: Cambridge University Press, 2008), 191.

8 27:365. In Collins' lecture notes the discussion of the duty of self-observation is described as a duty to observe ourselves through actions and to pay attention to our actions.



about the quality of my moral disposition are inferential and fallible, not introspectively certain.<sup>9</sup> Moreover, Kant is clear that there are limits to self-observation (6:63): “even a human being’s inner experience of himself does not allow him so to fathom the depths of his heart as to be able to attain, through self-observation, an entirely reliable cognition of the basis of the maxims which he professes, and of their purity and stability.” However, in insisting on the limits of self-observation Kant does not claim that self-observation is entirely unreliable.

There clearly is a point to Kant’s claim that we should judge ourselves and others, to the extent we need to, by our actions and not our “pious ejaculations” (27:365). But all (or most) judgment-corrupting influences that apply to introspective analysis hold also for the judgment of deeds. If self-conceit disposes me to a flattering self-description of my motives on a particular occasion, then surely it similarly complicates objectivity when my deeds are the object of analysis. Did I tell a lie or did I spare her feelings? Did I act out of a concern for her well-being, or did I act on a desire to avoid social unpleasantness? Did I give him his moral due, or did I try to leave him no recourse for complaint? Because actions are individuated by their maxims, it is not at all clear how the appeal to deeds can help to secure self-knowledge. This issue is problematically undermined by Kant’s exaggerated suspicion of the ineliminable motives of self-conceit.

Kant’s rejection of any purely introspective certainty fits with his exaggerated claims about its dangers. Self-observation, described as a “methodological compilation of the perceptions formed in us” (7:134), can “easily lead to enthusiasm and madness.” Kant is perhaps at his most non-Buddhist when he warns of the dangers of spying on oneself (7:134) and of “the affected composition of an inner history of the *involuntary* course of one’s thoughts and feelings.” His warning is grounded in his conviction that spying on oneself leads to “illumism or even terrorism” as it falsely leads one to believe in “supposed higher inspirations and powers flowing into us, without our help, who knows from where.” Kant goes on to mention several people, including Pascal, whom he takes to have been victim of this form of delusion. Clearly Kant has some very particular form of self-observation in mind in making such strong and somewhat peculiar claims. There is no doubt that

---

9 Owen Ware (16) argues that self-knowledge is inferential (not introspective) and that conduct evaluated or assessed is to be thought of in terms of “the verdict the agent could hope to receive if her *whole life* was placed before a judge.”

self-observation can be overdone, but Kant's apparent equation of self-observation with the practices of religious fanatics is crucially in need of an argument. One can accept Kant's worries about moral enthusiasm and the corrupting power of some religious doctrines and narratives, accept that self-observation can be overdone, and accept that our capacities to know our own motives are limited, without being forced to the conclusion that attending carefully to the contents of one's mind will be harmful to oneself.

It seems that the real problems for Kant's position arise out of his inflated suspicion of self-conceit and his exaggerated worries about the dangers of self-observation. In this, and in his adoption of duty as the framework for accumulating self-knowledge, Kant's pragmatic anthropology (his impure ethics)<sup>10</sup> is not empirically well-grounded. By framing knowledge of one's heart in terms of one's moral perfection in relation to duty, Kant makes it unlikely that the tendencies of self-conceit he worries about will be revealed.<sup>11</sup> Duty, after all, is constantly under threat from inclination. It is reasonable to think that more reliable access to knowledge of the source of my actions as pure or impure will can be had by framing one's experience in a manner less likely to stimulate the "dear self" (4:408) of self-conceit. In the next section I suggest that the Buddhist tradition offers just such a frame.

### 3. Buddhism and Self-Knowledge

The most famous teachings of Buddhism are the Four Noble Truths: the fact of suffering, the cause of suffering, the cessation of suffering, and the path to the cessation of suffering. Suffering refers to pain, discontent, and distress and the common examples used to illuminate the concept include the suffering of birth, aging, illness, and death. Among the steps in the Eightfold Path leading to the cessation of suffer-

---

10 Robert B. Louden uses the phrase "impure ethics" to refer to the part of Kant's moral philosophy that is empirically grounded in claims about human nature. See his, *Kant's Impure Ethics* (New York: Oxford University Press, 2000).

11 Elsewhere I have argued that Kant's method for attaining self-knowledge relies too greatly on the notion that duty, unaided by the cultivation of the moral emotions, can serve as a filter for self-illusion. See my "Moral Self-Knowledge in Kantian Ethics", *Ethical Theory and Moral Practice* 12 (2009), 525–37.

ing, Right View<sup>12</sup> (i. e., the ability to understand in one's experience suffering, its origin, its cessation, and the way to its cessation) is particularly relevant to the present discussion as a way of framing one's experience. The Buddhist view, like Kant's, is wary of our capacity to find in our self-examination what we have put there but, unlike Kant, Buddhism does not give up on the possibility of self-knowledge via introspective awareness. Instead, it offers a theoretical framework for experience explicitly intended to filter out self-illusion: the device used to defeat self-deception is the awareness of suffering.

But before proceeding to consider the Buddhist position, an objection to the suggestion that Buddhism can be said to promote self-knowledge must be addressed: how can self-deception be defeated if, as it is often claimed, Buddhism holds that there is no self, that it is illusory, or unreal? Within Buddhism the Not-self doctrine (*anatta*) is itself controversial.<sup>13</sup> Because self-knowledge, as awareness of both the nature of the mind and one's particular mind, is an essential part of the meditative practice definitive of Buddhist practice, it is important to see how the theoretical issue squares with the practical activity of acquiring self-knowledge. Hence, we must ask whether this doctrine is the assertion of a metaphysical truth (and so a bit of theoretical knowledge), as the name typically given to it, the "No-self doctrine", would imply, or a practical claim concerning a method of framing one's experience that allows one better to perceive and release oneself from suffering. The latter position is both more plausible and helps to shed light on some of the difficulties we encountered in Kant's account of self-knowledge. Thanissaro Bhikkhu has argued that the Not-self doctrine is best understood as a practical epistemic stance, a "technique of perception", and not as an ontological fact.<sup>14</sup> On this view, using Not-self as a technique of perception, one attends to one's experience and intentions in such a way that the sense of oneself as a discreet, constant and wholly independent being is diminished. *Anatta*, Thanissaro Bhikkhu argues, is better under-

---

12 "Right View" (*Sammaditthi Sutta*), in *The Middle Length Discourses of the Buddha: A New Translation of the Majjhima Nikaya*, trs. Bhikkhu Nanamoli and Bhikkhu Bodhi (Boston: Wisdom Publications, 1995), discourse 9.

13 For a brief but clear explanation of some of the metaphysical issues relevant to the nature of the Buddhist self, see Mark Siderits' *Buddhism as Philosophy: An Introduction* (Indianapolis: Hackett Publishing Company, 2007).

14 Thanissaro Bhikkhu, "The Not-self Strategy", Access to Insight website, <http://www.accesstoinsight.org/lib/authors/thanissaro/notself.html> (accessed July 6, 2009).

stood as a commitment, theoretical and practical, to apprehension in terms of “Not-self” rather than as a doctrinal commitment to a metaphysical fact of “No-self”.<sup>15</sup>

Making reference to important passages in the Pali Canon, Thanissaro Bhikkhu makes the case for Not-self as a technique of perception, noting that in the sole Sutta where the Buddha is directly asked to take an ontological stand on Not-self, he refuses to answer the question, explaining that, were he to answer, he would confuse and mislead.<sup>16</sup> Furthermore, in other Suttas the Buddha ranks the Not-self doctrine among those views that constitute a “fetter of views” and a potential impediment to right view insofar as it is, when adopted as a full belief about the nature of reality, itself an unskillful act. Because a commitment to an ontological version of Not-Self can lead one to focus attention on ideas that leave one bound to stress (unable to identify stress and its cause), it seems suitable to adopt it as a perceptual strategy rather than as a fact about reality, an ontological (No-self) commitment to be defended. The relevance of Not-self lies in its role in bringing an end to suffering, not in its ontological propriety, and the point of this form of mental regulation is not to express a recognition of a metaphysical fact about the universe. One avoids the fetter of views that talk of self and not-self leads to, in part, by apprehending experience in the following way: “He attends wisely: ‘This is suffering’; he attends wisely: ‘This is the origin of suffering’; he attends wisely: ‘This is the cessation of suffering’; he attends wisely: ‘This is the way leading to the cessation of suffering.’”<sup>17</sup> Thus we can see that the doctrine of Not-self does not ground an objection (based on the claim that there is no self) to Buddhism as a defender of self-knowledge. As a practical stance, Not-self actually serves as a means of promoting self-knowledge.

In any case, the practical or epistemic stance has priority in the present discussion if only because it is required for appropriate attention (*yo-*

15 This issue is discussed in “No-self or Not-self?”, in *Noble Strategy* (Valley Center, Ca.: Metta Forest Monastery Publisher, 1999), 71–4. I will follow Thanissaro Bhikkhu in describing *anatta* as Not-self, for reasons that will become clear in what follows.

16 “Ananda (Is There a Self?)”, *The Connected Discourses of the Buddha: A New Translation of the Samyutta Nikaya*, vol.II, tr. Bhikkhu Bodhi (Oxford: Pali Text Society, 2000), 44.10.

17 “All the Taints”, *The Middle Length Discourses of the Buddha: A New Translation of the Majjhima Nikaya*, trs. Bhikkhu Nanamoli and Bhikkhu Bodhi (Boston: Wisdom Publications, 1995), (Sabbasava Sutta) 2.

*niso manasikara*), itself a crucial factor in the development of skillful dispositions. Appropriate (or wise) attention refers to the ability to frame the issue of suffering in such way (in any circumstance) that one may deal with it skillfully, and is identified as extremely important in transformative practice.<sup>18</sup> Insofar as the Not-self doctrine, understood as a technique of perception rather than as an ontological commitment, supports appropriate attention, it seems the better interpretation. On this account, appropriate attention is best distinguished from inappropriate attention by its capacity to identify directly the most relevant aspects of immediate experience and to do so without framing them in terms of abstract categories. As Kant himself recognized, thinking driven by abstract categories such as “the noble” or “the magnanimous” is open to corruption by hidden motives and agendas that lie behind the concepts we commit to. It is best not to consider what one is doing in terms of its nobility, or magnanimity, as such honorifics are likely to activate the tendency to self-conceit that distorts judgment. Attempts at honest self-evaluation will be impeded by such notions. Kant recognizes the need to avoid such impediments to honest self-scrutiny and this is why he advises that we use the requirements of duty as our gauge. But although duty is in conflict with self-conceit, it is exactly the kind of abstract honorific that can be corrupted by the hidden motives it opposes. The image of duty can cast a glamour. So while for Kant, we are to develop self-knowledge within the framework of attention to duty, for Buddhism, attention to suffering is the framework, and because suffering is less liable to corruption by self-conceit, it is a better frame for honest self-scrutiny.

According to Buddhism, the common element in all forms of mental suffering is clinging or attachment to the five aggregates of form, feeling, perception, mental fabrication, and consciousness. So appropriate attention, as it pertains to the issue of self, is attention directed at one's experience in such a way that one will be able to find in that experience the elements of self that must be seen and comprehended in order for the self to be transformed. Specifically, by employing Not-self one comes to see that the contents of one's consciousness are not themselves constitutive of the self. Doing this helps to refine one's sense of self through a recognition of how it is that mental formations are fabricated. Kant's version of appropriate attention, by contrast, has

---

18 “The Paying Methodical Attention Sutta”, *The Itivuttaka*, tr. Peter Masefield (Oxford: The Pali Text Society, 2001), (*Yonisomanasikarasutta*) 1.16, I.9.

us look always to Duty. Suffering or stress (*dukkha*), on the other hand, is directly available in the contents of immediate experience and so the focus on suffering utilizes a phenomenologically trustworthy vehicle of knowledge.

As one develops more insight into the contents of one's mind, one comes to be better able to see the drawbacks of certain forms of activity and their role in suffering. This growth in comprehension plays an important role in self-governance: autonomy is enhanced by a clearer understanding of one's motives and the causal factors conditioning them. I can recognize over time, for example, that I am suffering because I fear the disdain of my colleagues; I can then proceed to challenge or abandon this fear. Self-knowledge framed by right view promotes autonomy and morally good action. Insofar as self-knowledge reveals not merely my own condition but that of humanity in general, it breaks down the barrier between self and other. To understand that *my* suffering is caused by *my* clinging, say, to the approval of others, is to understand a form of human suffering also experienced by others, also caused by clinging. Suffering is the great equalizer and, when appropriately framed, it can become the trustworthy bit of experience necessary for self-improvement.

#### 4. Kant and Buddha on Self-Knowledge

My point in this brief contrast between Kant and Buddha is to draw attention to several important similarities in their views, to highlight some of Kant's insights, and make some suggestive remarks about where Kant's account of self-knowledge goes wrong. Both think self-knowledge is crucial to good moral conduct and to self-transformation. Both think self-knowledge must be regulated through a theoretical frame, and both recognize this frame is necessary to exclude forms of attachment to the self that impede objectivity. The Buddha, who would include self-loathing as a form of attachment to the self, casts the broader net, not assuming, as Kant does, that duty is constantly at risk of subversion by inclination and self-conceit. Kant's supposition that duty is the appropriate frame for self-knowledge is the result of his conclusion that it is the only thing of any genuine value in us: as psychological creatures we can have no value in comparison with the sublime value of practical reason.

Kant fails to recognize what can be gained through introspective awareness and never provides a convincing case against what he derogatorily refers to as “self-observation”, but instead uses it to describe the moral enthusiasm that he rightly opposes. When he writes that “to wish to play the spy upon one’s self, when those acts come to mind unsummoned and of their own accord ... is to reverse the natural order of the cognitive powers, since then the rational elements do not take the lead (as they should) but instead follow behind” (7:133), he seems *again* to be worrying only about religious enthusiasts, and because he overestimates the transformative powers of duty as a theoretical frame for experience, he fails to recognize the resources that self-deceived beings like us have for self-knowledge. Buddhism, of course, finds spying upon oneself very useful.<sup>19</sup>

The practical counterpart of Kant’s elevation of practical reason to absolute value, as against my own moral self-awareness, is a kind of paranoia about the self as a deceiver and defeater of objectivity. Only the majesty of the moral law can keep us from the sly self-regard of the psychological self. Kant’s practical anthropology is somewhat contaminated by an *a prioristic* view of the self. His accounts of, and concerns about, moral enthusiasm and self-observation make evident the ways his theory of human nature remains problematically rationalistic. In contrast, the focus in Buddhism lies within experience, on the mundane facts of suffering. The relevant forms of self-knowledge needed to purify the will are not ultimately inaccessible but can be developed through an analysis of mental formations that can be refined and eliminated. Its methods will be more effective in achieving the purification of the will Kant expects of self-knowledge, than will a comparison of one’s moral perfection in relation to duty. In conclusion, although Kant’s insight into the moral need for self-knowledge is admirable, Buddhist moral psychology offers a vista of rich theoretical resources to develop the kind of awareness needed to diminish self-deception and promote right action.

---

19 This, of course, is not a careful description of the sort of self-awareness Buddhism promotes; it is clearly not self-directed espionage. However the reference to spying is apt insofar as we do secret away information that we do not wish to be uncovered.

## 54. Kant and Vasubandhu on the “Transcendent Self”

*Soraj Hongladarom*

In the Transcendental Deduction Kant argues for the existence of the “transcendental unity of apperception”, functioning as a principle of unity that makes synthesis of the manifold possible; synthesis in turn makes it possible for one to be able to claim with justification that one has a right to objective knowledge. What is really intriguing in Kant’s argument here is how the transcendental unity of apperception, or in other words transcendental self-consciousness, is different from ordinary empirical self-consciousness. Another related question is how the consideration of the two unities of apperception here is related to the whole set of questions pertaining to the self and the person. If the empirical unity of apperception is closely related to the ordinary conception of the empirical, individual self, then what does the transcendental unity of apperception correspond to?

I call the kind of putative self that corresponds to the idea of the transcendental unity of apperception, the “transcendent” self. The idea is that the transcendent self appears to exist over and above an empirical self, and functions as the condition of the possibility of the latter. But then a number of vexing questions arise. First of all, what exactly is the transcendent self? If the transcendental unity of apperception is to be able to do any real work at all, it has to possess at least some kind of status as something that is referred to when talked about. Then the question is exactly *what* is being talked about here? Kant’s text on this is, as is well known, exceedingly obscure; nonetheless it is my hope, and indeed my main contention in the essay, that one might understand this better philosophically if one considers the whole issue in light of insights obtained through comparison with the work of another philosopher from a very different tradition. I argue that the main work of a fourth century AD Indian philosopher and Buddhist saint Vasubandhu, the *Vimśatikā* (*Twenty Verses*), could shed light on this very difficult topic.

The idea is that both Kant and Vasubandhu make use of the distinction between the empirical self and the transcendent one. (My use of “transcendent” here only means that the status of the self in question is emphasized as some entity that at least functions as the referent of



phrases such as “the transcendental unity of apperception” and the like.) In what follows I shall discuss Kant’s main argument as well as Vasubandhu’s to the extent that is possible within the limited scope of this essay. Then I shall outline how Vasubandhu’s viewpoint could contribute to the problem of the status of the transcendent self as well as its philosophical functions.

Let us look again at Kant’s famous passage on the necessity of the “I think” to accompany all my representations (B131–2):

It must be possible for the “I think” to accompany all my representations; for otherwise something would be represented in me which could not be thought at all, and that is equivalent to saying that the representation would be impossible, or at least would be nothing to me.<sup>1</sup>

The idea is that, no matter what kind of representations I am having, I need to be able to put “I think” in front so that they are in fact *my* representations. It might be possible for me to entertain in some sense a representation I am not conscious of, but then that thought or representation would, in Kant’s sense, be nothing for me because it will not be possible for that representation to fall under the conceptual scheme that operates through the pure concepts of understanding. Even if it were possible that I have a thought I am not conscious of, that thought would then fall entirely outside my scope of understanding. This is Kant’s first step in arguing that all of my representations need to be able to fall under the scheme of the pure concepts of understanding. Furthermore, even if I could be conscious of one representation at one time and another representation at another time, if I could not relate these representations to fall under the same scheme, so to speak, then in a real sense neither of these representations could be called mine.

Kant then continues (B132):

That representation which can be given prior to all thought is entitled intuition. All the manifold of intuition has, therefore, a necessary relation to the “I think” in the same subject in which this manifold is found. But this representation is an act of *spontaneity*, that is, it cannot be regarded as belonging to sensibility. I call it *pure apperception*, to distinguish it from empirical apperception, or, again, *original apperception*, because it is that self-consciousness which, while generating the representation “*I think*” (a representation which must be capable of accompanying all other representations, and which in all consciousness is one and the same), cannot itself be accompanied by any further representation. The unity of this apperception I like-

---

<sup>1</sup> Kant, *Critique of Pure Reason*, tr. Norman Kemp Smith (London: St. Martin’s, 1929).

wise entitle the transcendental unity of self-consciousness, in order to indicate the possibility of a priori knowledge arising from it.

The main difference between what Kant calls “pure apperception” and “empirical apperception” is that the former, being the condition of the possibility of the latter, in fact originates from the understanding; it is pure apperception, or the “transcendental unity of self-consciousness”, that gives rise to the empirical representation to myself that it is my consciousness that accompanies my representations. In other words, the transcendental unity of self-consciousness is the source of the empirical I that accompanies my representations. Kant here employs his usual move of arguing for the necessity of the *a priori* as the condition of the possibility of what is already there empirically.

So the question is: How is one to understand Kant’s notion that the transcendental unity of self-consciousness “gives rise” to “the empirical representation to myself that it is my consciousness that accompanies my representations”? When I am conscious of myself, such as when I think of myself typing out this essay on the computer, it appears that my “self” exists at two levels. On the one hand, there is clearly somebody who is typing on the keyboard at this moment, and on the other there seems to be a rather different one who is being conscious of the act of typing. When we think about what we think, feel, desire, and so forth, what we think of are episodes of mental acts, and if these episodes were not threaded together within a single framework of self consciousness, then Kant would say that the episodes would not be mine at all. A consequence, of course, will be that no objective knowledge is possible since objective knowledge is possible only if one is able to thread the various episodes of one’s mental acts together under a single framework. Indeed this is the conclusion of the Transcendental Deduction itself. So it is quite clear that the episodes themselves are one thing, and the act of synthesis that presupposes the transcendental unity of self consciousness is another. Nonetheless, any object of self consciousness, any referent of the first person pronoun in thoughts of the type “I am thinking that F” needs to be empirical because it is being thought of. For Kant, the transcendental unity of apperception functions solely as the source of the possibility of awareness of empirical episodes of the various selves acting in various moments: “it is that self-consciousness which, while generating the representation ‘I think’ (a representation which must be capable of accompanying all other representations, and which in all consciousness is one and the same), cannot itself be accompanied by any further

representation.” It cannot be accompanied by any further representation because it is the source, the condition of the possibility, of there being empirical awareness of the self doing such and such from the beginning.

The idea here bears similarity to what Vasubandhu says in his *Vimśatikakārikā*. What ties the two works together is that they are both idealist. Kant’s, of course, is transcendental idealism, meaning that it is an idealism about the thing in itself and not about ordinary empirical objects in general. Vasubandhu belongs to a philosophical school in Buddhism known as “Yogācāra”, espousing that reality as perceived is nothing but consciousness. That is, ordinary objects are ultimately speaking nothing but projections of the consciousness, deriving their being and characteristics from consciousness. It is not a straightforward matter, however, whether Yogācāra should be classified as an empirical or transcendental idealist, since transcendental idealism is Kant’s own terminology and presupposes his own philosophical system. Nonetheless, this matter does not have to concern us here, since we are considering the two levels of the self and their philosophical implications. For this matter, let us look at a part of the seminal text of this topic, the *Vimśatikakārikā*:

... we must distinguish between reality [self and objects] as constructed by ordinary consciousness (especially the imagination) and reality as it is in itself, in its “suchness” (*tathatā*). Beyond the ordinary (constructed) self [ego] and its subject–object duality, there is an ineffable (*anābhilāpya*) transcendent Self (in which the duality of subject and object does not arise), which is known by the Buddha and other enlightened ones. It is the constructed self and its constructed objects that are insubstantial, merely transformations and representations of consciousness .... [The ineffable (true) Self is substantial (*dravyatah*), that is, “really real.”]<sup>2</sup>

For those who are not familiar with Buddhist philosophy, this can present a real challenge. Nonetheless, our purpose here is more modest: to find similarities or differences between what Vasubandhu is saying here and Kant’s view on the unity of apperception. What Vasubandhu is saying here is that there is a distinction between the self as object of perception and the “transcendent” Self that is ineffable and beyond the duality of subject and object. In the passage Vasubandhu does not present an explicit argument for the existence of the transcendent Self, but it can

2 Vasubandhu, *Vimśatikakārikā* (“Twenty Verses on Consciousness Only”), tr. George Cronk. Available at <http://www.bergen.edu/phr/121/vasubandhugc.pdf> (retrieved May 17, 2009), 171.

be inferred that such a self needs to exist because if it did not, there would then be no grounding of unity of an individual's self in such a way that coherence in thought and understanding is possible. What is startling here is Vasubandhu's assertion that the transcendent Self is ineffable and beyond subject-object duality. This requires at least some explication. What Vasubandhu seems to have in mind is that what is effable, that is capable of being expressed through language, requires that there be a distinction between subject and object. In other words, the duality or distinction between subject and object is necessary for formation of an expression in propositional form so that a judgment is possible. Any proposition must consist of a subject and a predicate, and this distinction seems to mirror the distinction between subject and object in the mind. One talks about one thing, ascribing certain qualities to it; hence there is a distinction between the thing talked about and the qualities ascribed.

Kant terms this distinction as one between intuition, as the matter of judgment, and concepts, corresponding to the predicate in the form of judgment. In terms of the awareness of the self, in so far as the self here exists as an object of thought, it is effable because it exists within the scheme of the duality of judgmental form. This is but another way of saying that the empirical self is always effable. However, when it comes to Kant's transcendental unity of self-consciousness, a question then emerges, whether it can be only a subject. A related question, whether the transcendental unity of self-consciousness can be an object, can be answered in the negative from the beginning because it functions as the origin of *a priori* synthesis; hence it cannot fall under the scheme whereby it is being thought. However, if the transcendental unity of apperception (or self-consciousness) is only a subject, then it always requires an object. Furthermore, since there does not seem to be a guarantee that a subject accompanying an object needs to be one and the same, the transcendental unity here cannot perhaps be a subject either. So, if we accept the argument that a necessary condition for effability is that it falls under the subject-object distinction, then the transcendental unity of self-consciousness here does not appear to be effable.

That the transcendental unity of self-consciousness is ineffable does not mean that it is mystical; rather it means that it underlies the very possibility of effability and of any objective relation between the pure concepts of understanding and empirical intuitions. Here the transcendental unity of apperception functions in a similar way to Vasubandhu's transcendent Self in that it grounds the very coherence of thought that

alone makes possible a person's objective, empirical knowledge. In this sense both the transcendent Self and the transcendental unity of apperception do not belong to one person only. It is not the case that my "transcendental unity of apperception" and yours are numerically distinct or numerically one and the same, since the very concept of identity, being one of the pure concepts of understanding, already presupposes the transcendental unity. Thus identity does not apply to the latter.

# 55. Kant's Moral Philosophy in Relation to Indian Moral Philosophy as Depicted in *Srimad-Bhagavad-Gita*

*Ruchira Majumdar*

## 1. Introduction

The essay discusses Kant's ideas and views on moral philosophy from the perspective of Indian moral philosophy, especially that of *Srimad-Bhagavad-Gita*. The essay portrays the great German philosopher Immanuel Kant as an ascetic who could comprehend the same moral insight as that of Lord Krishna, the epitome of Indian divinity and author of the greatest sacred holy book of India, the *Srimad-Bhagavad-Gita*.<sup>1</sup> The name "Srimad-Bhagavad-Gita", when translated into English literally, means "Song from (the mouth of) God". In this book Lord Krishna explains the nuances of ethico-religious life to his disciple Arjuna, who expresses unwillingness to participate in the holy war at Kuruksetra since the opponents include his own cousin, brothers, grandparents, other relatives, and teachers. He prefers to leave his kingdom and survive by alms to killing his own cousin, brothers, and relatives. Overwhelmed by grief, he seeks his friend Krishna's moral advice. At this point in time, Lord Krishna explains how Arjuna should use his reason, determine his duties, and perform them without any desire or concern for consequences (*niskama karma*). He has to fight evil and protect the innocent people as a king, because this is a king's duty.

Kant's moral deliberations share many common concepts with the teachings of *Srimad-Bhagavad-Gita*: both indicate how to attain *personhood* by removing the barriers of individuality that arise out of our animalistic faculties, desire, and passion, giving rise to actions done under desire for the end (*sakama karma*). Both Kant and the divine prophet,

---

1 For references to *Srimad Bhagavad-Gita*, see S. Radhakrishnan, *The Bhagavad Gita* (Hollywood: Vedanta Press, 2009), and Sri Aurobindo, *Essays on the Gita* (Pondicherry, India: Sri Aurobindo Ashram Press, 1995).

Lord Krishna (who is also identified as Jagannatha, Balaji, Tirupatiji, etc., in different incarnations), stress that our rational faculty should be used as the moral torch to determine one's duty and to help us transcend the sensuous limits of our physical existence. Thus, attainment of *human personhood* is the sole aim of both thinkers; because of these and many other similarities, we think it is wholly appropriate to trace a comparative discussion between them. The essay recapitulates the main theme of Kant's moral philosophy and that of *Srimad-Bhagavad-Gita*, then draws points of similarity between Kantian and *Srimad-Bhagavad-Gita's* moral philosophical tenets.

## 2. Major Conceptual Framework of Kant's Moral Philosophy

Kant's moral philosophy is based on several important themes. These are: (1) practical reason as a Rationalism through the moral law that leads mankind down the moral path; (2) the categorical imperative as the moral law that implores human beings to keep on moral path; (3) duty for duty's sake, or the performance of one's duties irrespective of the result, circumstances, and agent's attitude; (4) goodwill; (5) the postulates of morality; (6) maxims of morality as the general principle that provides guidelines for moral life; and (7) the complete good or virtue-cum-happiness, being the ultimate end of moral life.

To begin with, Kant has mentioned two principal faculties governing human life, sensibility and reason, that often pose conflicts in the moral path. As Seth points out: "As a sentient being, man is a member of the animal sphere, whose law is pleasure, as a rational being, he enacts upon himself the higher law of reason which takes no account of sensibility."<sup>2</sup> In his book *Fundamental Principles of Metaphysics of Morals*, Kant distinguished between "theoretical pure reason" and "practical reason" : while the former is used in epistemology and metaphysics, the latter provides us with moral instructions and prescriptions in everyday life.

In a Kantian sense the moral law is deduced from our practical reason, a universal and objective element in human life. It is contrary to our animalistic sensuous faculty that is individualistic and egoistic.

---

2 James Seth, *A Study of Ethical Principles* (Whitefish, MT: Kessinger Publishing, 2007).

Thus, while practical reason and the moral law guide us in the path of complete personhood, sensibility leads us to narrow individuality. Kant prescribes control of sensibility by the moral law for achieving *personhood*, as the epitome of moral goodness. In other words, a morally good human being is a person, who has attained virtue-cum-happiness. The themes in Kant's moral philosophy can be further developed as follows.

(1) *Rationalism*: Kant's moral philosophy is known as Rationalism, Rigorism, Moral Purism or Formal Ethics. It is opposed to Hedonism, emphasizing the claims of sensibility. Kant's rationalism highlights the claims of reason and advocates self-denial and self-conquest. It seeks to spiritualize natural human beings, differentiates us from the lower animals, and imbibes in us the life of pure reason. The fulfillment of the higher self or pure reason is the ideal of Kant's rationalism.

(2) *Moral law*: Kant's moral philosophy is akin to Rational Intuitionism. Kant regards conscience as practical reason imposing the moral law upon itself. The moral law is deduced from practical reason and known intuitively along with its maxims. The moral quality of an action is not determined by any end or consequence but by its agreement or disagreement with the moral law. Thus Kant is an advocate of deontological or jural ethics as distinguished from teleological ethics.

(3) *Categorical imperative*: The moral law is a categorical imperative and the ultimate moral standard. It is an imperative or command as opposed to a description. It is a categorical as opposed to a hypothetical or assertorical command. It is unconditional, known *a priori*, and its imposition does not depend on any subjective factors like consequences, circumstances, and the agent's attitude.

(4) *Duty for duty's sake*: As a consequence of the categorical imperative, one's duties are deliberately performed by an agent, not for any other consideration, but only for the fact that they are duties. For example, Kant avoids even factors like love, care, attention, etc., as the guideline for helping others. He explains that, if one helps another only due to these factors, then the duty of helping others becomes relative to these personal factors, and loses its objectivity and universality. A mother should take care of her child not because of her affection, but because it is her duty to do so. Thus, Kant ensures the care and protection of even unwanted children, old parents, and other weaker members of the society.

(5) *Goodwill*: Kant's first principle is that "there is nothing in world or even out of it that can be called good without qualification except a



good will”.<sup>3</sup> It is the only intrinsic value, while all other values are extrinsic. It, “like a jewel, would still shine by its own light”,<sup>4</sup> even if it fails to achieve anything. The good moral action is an action done from good will, and the essence of a good action is the good will that inspires the agent to perform the action.

(6) *Postulates of morality*: Kant admits three postulates of morality. These are the immortality of soul, freedom of will and the existence of God. An action done by a free will is a moral action. The rational soul is immortal and God rewards it with happiness for the virtues accumulated by it, even in the next world. God is the controller of the realms of nature and spirits. These three postulates can be very briefly explained as follows.

(a) *Freedom of will*: The human being is essentially a self-determined and free person. Kant believes in noumenal freedom and empirical necessity and advocates indeterminism. He argues that in moral judgments there is a sense of “ought” or moral obligation that implies free will. The maxim “Thou oughtest, therefore thou canst” encapsulates this truth.<sup>5</sup> Duty, responsibility justice, accountability, merit and demerit, virtue and vice, praise and blame, reward and punishment, would be meaningless if the agent does not have freewill. Denial of freewill saps the very foundation of morality and cuts at the root of responsibility.

(b) *Immortality of the soul*: The moral ideal is infinite and cannot be fully realized in the present finite life. The realization of the moral ideal, therefore, requires an immortal life. Kant argues that the conflict between desire and duty cannot be reconciled in the finite life. So there should be future lives continuous with the present life, where the personality of the human soul will persist and the conflict between desire and duty will be completely removed. I call this demand of our moral nature “vaccination of the conscience”. Kant also argues that the demand of the conscience for justice and equity points to the existence of future lives. Virtuous persons are seldom found to be happy, so there should be after-lives when God will control the perfect harmony of virtue with happiness.

---

3 Immanuel Kant, *Fundamental Principles of Metaphysic of Morals*, tr. T. K. Abbott, in *Immanuel Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*, fifth edition (London: Longmans, Green & Co., 1898), 9 (section I).

4 Kant, *Fundamental Principles*, 10 (section I).

5 See W. Lillie, *An Introduction to Ethics* (New Delhi: Allied Publishers, 1957), ch. VIII.

(c) *Existence of God*: God is the eternally perfect reason who has supreme power and, as the moral governor, rewards virtuous agents. Our moral ideal is ideal to us; it is real to God.

(7) *Principles of morality*: Kant tries to make the moral law more definite by laying down a number of universal maxims. These are:

(a) “Act only on that maxim which thou canst will to be a universal law”;<sup>6</sup> this explains that what is right is universal, and that what is expedient is not.

(b) “So act as to treat humanity whether in their own person or in that of any other, always as an end, and never as a means only”; this requires us to treat personality as of an absolute worth. This maxim has a corollary: “Try always to perfect thyself, and try to conduce to the happiness of others, by bringing about favourable circumstances, as you cannot make others perfect”;<sup>7</sup> this stresses the fact that moral will has to be cultivated by each agent individually.

(c) “Act as a member of the Kingdom of ends”; this advises us to treat ourselves and others as of equal intrinsic value, regarding all persons as ends in themselves. In an ideal society, governed by the moral law, each person realizes his or her own good by promoting the good of others.

(8) *Complete good or “virtue-cum-happiness”*: God will harmonize virtue with happiness and bring about the complete good of the performer of good actions. For, our moral sense demands that the complete good cannot be less than something mingled with happiness and bliss.

### 3. Major Conceptual Framework of *Srimad-Bhagavad-Gita*

*Srimad-Bhagavad-Gita* is an important Sanskrit Hindu scripture and is revered as a sacred text of Indian ethico-religions theory. It comprises 700 verses, originating from the famous epic *Mahabharata*. The prophet of this scripture is Lord Krishna, who is regarded as the supreme manifestation of divinity. He is the epitome of *personhood* or *purusottama*.

As mentioned at the outset of this essay, the content of *Srimad-Bhagavad-Gita* is a conversation between Lord Krishna and Arjuna, taking place on the battlefield before the start of the Kuruksetra war. Respond-

6 Kant, *Fundamental Principles*, section II.

7 *The Metaphysical Elements of Ethics*, iv-vi (Preface), tr. Abbott, 296–302.

ing to Arjuna's confusion, devastation, and moral dilemmas, Krishna explains Arjuna's duties as a warrior and prince, and elaborates on different Yogic and Vedantic philosophical themes, with examples and analogues. The basic themes of *Srimad-Bhagavad-Gita* are as follows.

(1) Reason (*buddhiyoga*) should guide our actions (*karmayoga*) and we should perform them with dedication (*bhaktiyoga*). This view is called *janakarmabhakti samuccayavada*, where reason, action, and dedication lead persons to their moral end.

(2) *Svadharna* is one's specific duty while *sadharana dharma* is universal duty that is applicable to all. Both are to be performed by a rational being, irrespective of the circumstances, consequences, and agent's attitude.

(3) *Karmayoga*: One has to perform one's duty without any consideration for the end. These are *niskama karma* as opposed to *sakama karma*, actions done for the sake of some result. Distinctions have also been made between *karma* or proper actions and *vikarma* or wrong actions and *akarma* or inaction. In *Srimad-Bhagavad-Gita* Krishna prescribes performance of *svadharna* without the motivation of result. This is the *karmayoga* of *Srimad-Bhagavad-Gita*.

(4) *Control of sense-organs*: Lord Krishna compares the sense-organs with poisonous snakes that vitiate a person with worldly desires and passions, limiting one within one's own individuality. He prescribes *buddhiyoga* to control sensibility and uplift morality.

(5) *Immortality of the soul*: The soul is eternal; it is without birth and death, pure and without form. Just as a person discards old clothes and wears new ones, similarly with the soul, a person discards an old body and takes up a new one. This is the *transmigration (janmantara)* of the soul.

(6) *Salvation and union with God*: The ultimate end of life is *moksa* (salvation) from the misery and sorrow that arise due to attachment with body. After one attains salvation, one is identified with the divine Supreme Person (*Brahma*) and attains Brahmisthiti. One thereby becomes a complete person and a *sthitaprajna* (a person with permanent insight).

#### 4. Kant's Moral Philosophy vis-à-vis Indian Moral Philosophy as Depicted in *Srimad-Bhagavad-Gita*

Firstly, both Kantian moral law and *Srimad-Bhagavad-Gita* prescribe that a moral life is a life guided by reason. Kant coins the term “practical reason”, while *Srimad-Bhagavad-Gita* approves of “*Buddhiyoga*”. However, neither of them prescribes annihilation of sensibility, but opted for control of the sense-organs by reason.

Secondly, Kant introduces the notion of the “categorical imperative” as the moral law. It is an absolute, *a priori* imperative, deduced from practical reason, that we know intuitively. It is applicable to all rational beings irrespective of caste and creed, social and financial status. For example, “speak the truth” is a categorical imperative that is to be followed by any rational being who has even the potential to speak. In *Srimad-Bhagavad-Gita*, the ideal of actions done without desire is similar to the Kantian notion of the categorical imperative. As Lord Krishna dictates, whatever may be one's socio-economic standing, one must perform actions without any desire for the end. Thus, the ideal of “*Svadharmā*” is closely akin to the Kantian notion of the categorical imperative. It is stated that even one's death while performing one's *svadharmā* is better than not performing it (*svadharme nidhanam-reya*).

Thirdly, the Kantian notion of “duty for duty's sake” is absolutely identical with the notion of “*Niskama Karma*” or “*Karmayoga*”. Kant took up the mantle of deontological morality when he stated that one should be directed by only the moral law in identifying one's duties and performing them. There should be no consideration of circumstances, consequences, and the agent's attitude. For if we take these factors into consideration, the action would be relative and subjective and lose its moral flavor. *Srimad-Bhagavad-Gita* reiterates that “*Karmanyeva adhi-karaste ma phalesu Kadacana*” (i. e., “you have only *rights towards performing the action*, and never towards its consequences”).<sup>8</sup> It explains that one should get rid of all egoistic attitudes like “I am the performer”, “I am the enjoyer”, etc. One should not pay any heed to either antecedent or consequent circumstances. One should not be guided by any desire or passion, but only by the strong dictate of reason, that implores one to perform the duty.

---

8 *Srimad Bhagavad Gita*, section II, verse 28.

Fourthly, Kant states that the essence of a morally good action is the “goodwill” that results from practical reason. An action becomes morally good not because of its consequences, but only due to the goodwill that engenders it. It is the only intrinsic good, while all other goods are extrinsic. For example, intelligence becomes morally good only when it is used by good will, but it becomes evil when used by a bad will. In *Srimad-Bhagavad-Gita*, Lord Krishna stresses the *vyavasayatmika buddhi* or the self-determinate and uplifting knowledge that induces a person to perform an action for others’ good and social welfare. Only a person with *vyavasayatmika buddhi* is capable of performing such actions, and is called a *sthitaprajna* (i. e., one whose reason is absolute). Nothing can alter such a person’s mental stability, being indifferent to either happiness or sorrow. The *sthitaprajna* lives only for universal good and is the retainer of absolute goodwill or *vyavasayatmika buddhi*.

Fifthly, just as Kant has mentioned several “maxims” in order to explicate the moral law and its execution, similarly in *Srimad-Bhagavad-Gita* we find innumerable “maxims” relating to moral progress. For example, just as Kant highlights reverence for the moral law, *Srimad-Bhagavad-Gita* stresses reverence for one’s duty.

Sixthly, regarding the ultimate end of moral life, Kant introduces the notion of “virtue-cum-happiness” as the complete good. He associates non-sensuous rational happiness with moral virtue. He provides assurance for a common person that God is there to reward us with happiness when we reach the apex of moral life. In *Srimad-Bhagavad-Gita* Lord Krishna declares that anybody who performs the *niskama karma* throughout life would attain “*Brahimisthiti*” (i. e., union with God), a state of eternal bliss. Thus, in both views, the ultimate end is divine and blissful.

Seventhly, in the moral philosophies of Kant and *Srimad-Bhagavad-Gita*, we find a unique combination of *reason*, *action* and *faith* in God. Though Kant did not mention the faith factor at the beginning, he admitted it when explaining the notion of “complete good” and the “maxim” prescribing reverence for law. In *Srimad-Bhagavad-Gita*, however, these three factors are very prominent throughout the text, especially in *jnanakarmabhakti-samuccayavada*.

Eighthly, both Kant and *Srimad-Bhagavad-Gita* admit the immortality of soul and existence of God in order to explain the complete good of mankind. Kant explains that if a person, in spite of doing morally good actions, remains unhappy and miserable, God will provide the complete good of *virtue-cum-happiness* in the next world. Moreover, desire cannot be overcome in a finite life. Hence the immortal soul sur-

vives bodily deaths to eliminate finally one's desires through a number of lives. In *Srimad-Bhagavad-Gita* the soul is depicted as eternal, immutable, and spiritual. Its essence should be realized by means of meditation throughout different bodily lives. When all desires are eliminated by performance of *niskama karma*, the soul attains union with God, a state of eternal bliss.

Finally, Kant highlighted the notion of "personhood" when he mentioned the maxims: (1) treat every rational being including yourself always as an end and never as a means (*Metaphysics of Morals*); and (2) act as a member of the Kingdom of ends. In other words, we have to treat each human being as a person, and never as an individual whom we can use as a means for a greater end. Persons are ends in themselves, and never only a means. In *Srimad-Bhagavad-Gita* Lord Krishna resonates this feeling when he states that a person must be free from the fetters of the three elements: *sattva*, *rajas*, and *tamas*. He must attain *nistraigunya* and become an independent and complete person.

There are also many other aspects of similarity between the moral philosophies of Kant and *Srimad-Bhagavad-Gita*. Only a few prominent ones have been mentioned here.

## 5. Conclusion

The discussion here envisages that Kant's moral philosophy is akin to the basic moral theme of the Indian ethico-religious conceptual scheme. We may conclude this brief discussion with the observation that Kant had rare moral insight and purity of heart, and the basic instructions of *Srimad-Bhagavad-Gita* were revealed to him. This fact once again proves that the dictates of reason are universal and objective and identical throughout the ages, all over the world.

## 56. Human Personhood at the Interface between Moral Law and Cultural Values

*Takayuki Kisaka*

### 1. Universalizability and Material Values in Kant's Ethics

Kant's ethics is formal. But human personhood does not consist only of formal values in a logical sense. In Kant's formal moral law empirical-anthropological and Christian values are contained in some way. Concerning this point I shall distinguish in this essay between universalizability as a kind of logical principle and the material values in Kant's ethics.

Kant himself thought that pure practical reason is general and actually gives us the moral law in the "factum of reason" (5:31–2). The universalizability principle contained in it is surely general (i. e., valid all over the world) at least in a logical sense. But the material values contained in Kant's ethics are not always general. Kant postulates the immortality of our soul and existence of God, for example, are necessary for the possibility of the highest good. But in fact, we Japanese usually need neither the existence of God nor the immortality of soul in the Christian meaning, in order to be moral. For Kant himself, the concept of God based on that of the highest good is no less empty for being only practically helpful for morals (6:5):

... das ist, die Idee eines höchsten Guts in der Welt, zu dessen Möglichkeit wir ein höheres, moralisches, heiligstes und allvermögendes Wesen annehmen müssen, das allein beide Elemente desselben vereinigen kann; aber diese Idee ist (praktisch betrachtet) doch nicht leer: weil sie unserm natürlichen Bedürfnisse zu allem unserm Thun und Lassen im Ganzen genommen irgend einen Endzweck, der von der Vernunft gerechtfertigt werden kann, zu denken abhilft, welches sonst ein Hinderniß der moralischen Enschließung sein würde.

But Japanese could be also moral on the basis of another worldview. The material values and worldviews concerning moral human person-

hood can be very different according to other cultures. And our human personhood does not play an adequate role morally without such material values in any case. This indicates that our morally adequate human personhood requires material values, based on traditional culture and religion. As we will see, Kant's moral law consists in the application of the universalizability principle to empirical-anthropological and Christianity-oriented values.

## 2. The Universalizability Principle and Its Application to a Teleological Nature

Kant infers the prohibition of suicide, telling a lie, and laziness, and also the obligation of mutual assistance from his categorical imperative. We should distinguish his argument concerning telling a lie from the other three cases. The logical implication of generalizing the maxim itself contradicts the very aim of the maxim in that case. We need the prohibition against telling a lie in order to tell a lie successfully; our aim should not be for the lie to be known as such. On the other hand, the implication of generalizing the maxim does not simply contradict the maxim itself in the other cases. It seems the prohibition of telling a lie is also the constitutive principle of the pragmatic world in general. But in other cases we must rather investigate the results of generalizing the maxim in a certain given natural world. The point is whether the generalized maxim can be valid as a general natural law in the given world. Kant discusses this point as follows (4:421):

Weil die Allgemeinheit des Gesetzes, wornach Wirkungen geschehen, dasjenige ausmacht, was eigentlich Natur im allgemeinsten Verstande (der Form nach), d. i. das Dasein der Dinge, heißt, so fern es nach allgemeinen Gesetzen bestimmt ist, so könnte der allgemeine Imperativ der Pflicht auch so lauten: handle so, als ob die Maxime deiner Handlung durch deinen Willen zum allgemeinen Naturgesetze werden sollte.

In this formulation the viewpoint from nature is important. "Natur im allgemeinsten Verstande (der Form nach)" implies a teleological standpoint. According to this second criterion concerning nature, we can argue that the maxims allowing suicide, laziness, and disregard for the need of mutual assistance in difficulties cannot be valid, because they contradict the teleological and anthropological nature of our world.



### 3. The Teleological Order of Nature as an Empirical Fact Described in Metaphysical Terms that Inevitably have Cultural and Religious Backgrounds

The teleological and anthropological nature of our world may be empirically general (i. e., the consequence of the application of the universalizability principle could be robustly the same). But when we ask why morality exists, we usually answer this question in terms of metaphysical statements that appeal to our traditional cultural backgrounds. With these kinds of backgrounds specified, the universalizability principle could function more properly as a moral principle: the specific moral meanings would be understood better as the result of applying it to them. We can find the Christianity-oriented worldview behind the Kantian teleological and anthropological natural order. Kant finds the reason for the prohibition of suicide in the teleological nature of our senses, thus implying natural teleology (4:422). Natural teleology leads to moral teleology. And it leads to humanity as the final purpose of creation; in relation to this purpose he also postulates the immortality of our soul.<sup>1</sup> Kant's thought finally leads to practical and "dogmatic" metaphysics through the critique of reason.<sup>2</sup> Kant actually presupposes the culturally material metaphysics. The reason morality exists is the logical

- 
- 1 The postulate belongs to the sphere of moral discussion. The idea of God and that of immortality of our soul are only helpful in order to be moral. Moral principles are originally independent of such encouragement. But we do need some such kind of encouragement in order to be moral adequately. As Palmquist says, we can find "the elements of pure practical (= moral) religion" here. The actual kind of encouragement for Kant is naturally Christian. Palmquist develops very detailed discussions "in which Kant tests the applicability of his theory by using it to assess the rationality of Christian Religion" (Stephen Palmquist, *Kant's Critical Religion: Volume Two of Kant's System of Perspectives* (Aldershot, UK: Ashgate, 2000), 143–4, 189 f. But I think Kant's ethics is already Christianity-oriented, insofar as (especially) the postulate of the immortality of the soul also belongs to it.
  - 2 We find this terminology in Kant's 1791 essay, "*Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnizens und Wolfs Zeiten in Deutschland gemacht hat?*" (20:253–332). "Critical" and "dogmatic" are not contradictory in this terminology. Only after the critical assessment of any kind of metaphysical thinking can the metaphysics of morals be not only critically but also dogmatically valid in the scope of conscious relativism that is inevitably based on some culturally oriented worldview. This is the two-world theory in the case of Kant's worldview.

universalizability principle on one hand. But it needs to be applied to some material spheres, relating to descriptions of nature that have cultural and religious backgrounds. The metaphysics of morals is almost impossible without any cultural worldviews and the relevant ethical values. Kant's worldview is substantially oriented into the two-world theory. And as we will see, Kant is also obliged to make his own metaphysical orientation principle clear, being faced with the worldview of Spinozism. The metaphysics of morals needs at least some critical introduction of cultural and religious material. As we will also see, at this point we cannot avoid some degree of cultural relativism in the realm of metaphysics. But following Kant, moral metaphysics can be also compatible with the requirement that morals should be valid independent of cultural values.

#### 4. "The feeling (*das Gefühl*) of the right and left hands": the Principle of Cultural Value Embodiment

Kant introduces a subjective and cultural value orientation principle in his minor but very important work, "Was heißt: sich im Denken orientieren?" Its theme is how to deal with Spinoza's rationalism, but Spinoza's cultural "Other" is different from Kant's own worldview. According to Spinozism human thoughts are among the infinite variations of God's thoughts. Kant thinks this is not logically impossible. But he argues that we do not have any intuition to know whether it is objectively true. Everything that is beyond our sense is beyond our knowledge. Kant's metaphysical statements about teleology are also beyond our sense. Moreover, Kant admits that Spinozism is most rational if one thinks consistently without transcendental idealism, especially without the concept of the transcendental ideality of space and time (5:102): we cannot actually think about the creation of the spatiotemporal world and the transcendence of God consistently; compared with this ordinary line of thinking, Spinozism is consistent and more persuasive (*weit bündiger*).

Kant's criticism concerning Spinozism consists of two points. Firstly, the transcendental ideality of space and time must be a universal truth, but Spinozism is not based on it. Secondly, Spinozism without transcendental idealism would destroy God's transcendence, which lies at the core of the Christian worldview—in other words, at the core of a

view of human personhood that presupposes human freedom beneath the transcendence of God.

With regard to the second point concerning Kantian religious culture we should notice that Kant has already been rationally obliged to commit himself to cultural relativism. Let us suppose, so to speak, the generous kind of Spinozism that has a different value orientation principle than Kant's, but recognizes the truth of transcendental idealism and perhaps postulates the identity of some of God's thoughts and our thoughts. Nothing is then left for Kant to criticize this generous Spinozism. It is a rational Other that is only culturally different from Kant. This Other would have a different conception of happiness and morality from Kant's and would not have to postulate anything transcendent. But Kant and this Other would be able to understand each other's worldview-postulations within the truth of transcendental idealism. Transcendental idealism could thus play a role as a principle for the mutual understanding between cultural Others. Following Kant we should recognize the truth of transcendental idealism, then we would be rationally obliged to have our own value orientation principles in order to have an adequate view of human personhood. Kant actually has his own value orientation principle.

Kant introduces "das Gefühl des der Vernunft eigenen Bedürfnisses" as his own value orientation principle (8:134–6). He needs to orient his view of human personhood to the transcendent God. Its basis is the subjective function of reason to regard what is necessary for it to be true. This is also the principle of Kant's religion of reason. The basis of the principle is its rational function. But its positive core is "das Gefühl":

Zu diesem Behuf bedarf ich aber durchaus das Gefühl eines Unterschiedes an meinem eigenen Subjekt, nämlich der rechten und linken Hand. Ich nenne es ein Gefühl: weil diese zwei Seiten äußerlich in der Anschauung keinen merklichen Unterschied zeigen. ... Also orientierte ich mich geographisch bei allen objektiven Datis am Himmel doch nur durch einen subjektiven Unterscheidungsgrund ...

We cannot acquire the positive orientation principle by means of intuition. We rather need some real and bodily feeling. It is culturally self-evident on one hand, but cannot be justified by reason on the other hand. This feeling is the necessary counterpart of universal reason. It is the very Other of universal pure reason. Kant, who abandoned the objective justification of the two-world theory in *Grundlegung zur Metaphysik der Sitten*, expressing his standpoint with the phrases below, is

then obliged to introduce the principle of feeling by means of a physical analogy of bodily orientation (4:463):

Und so begreifen wir zwar nicht die praktische unbedingte Notwendigkeit des moralischen Imperativs, wir begreifen aber doch seine Unbegreiflichkeit, welches alles ist, was billigermaßen von einer Philosophie, die bis zur Grenze der menschlichen Vernunft in Prinzipien strebt, gefordert werden kann.

On the basis of this principle, Kant secures his own material value in the realm of logically possible worldviews in general. He can locate his cultural and individual belief in the world rationally. But this has nothing to do with either theoretical objectivity or political positivity. It has rather to do with, so to speak, intercultural publicity. Let us call this feeling-principle of thinking-orientation *the principle of cultural value embodiment*.

## 5. Kant is Very Conscious of His Own Position

Critical reason that is culturally embodied also admits culturally other sorts of rational belief within the truth of transcendental idealism. The objective world is something ideal, because the mathematically measurable physical world depends on our spatiotemporal form of sensibility. But the rationally subjective principle of cultural value embodiment is something real.

Kant's view of human personhood morally postulates God's existence and the immortality of our souls on the basis of the necessity of the highest good. The categorical imperative demands that we realize the agreement of morality and happiness. The agreement may be logically necessary for every view of human personhood. But we should distinguish this logical necessity from the material necessity of God's existence and the immortality of our souls. In many cultures, including Japanese culture at least, people think they can be adequately moral without such postulations. Moreover, we can find multiple conceptions of happiness all over the world. The Kantian conception is only a particular one, and his theory of postulation depends on his particular conception. But Kant is very conscious of this point. According to the second *Critique* our theoretical reason does not know how to think about the realization of the highest good and so leaves the problem to the wisdom of practical reason (5:144–5):

Was aber das zweite Stück jenes Objects, nämlich die jener Würdigkeit durchgängig angemessene Glückseligkeit, betrifft, so ist zwar die Möglichkeit derselben überhaupt einzuräumen gar nicht eines Gebots bedürftig, denn die theoretische Vernunft hat selbst nichts dawider: nur die Art, wie wir uns eine solche Harmonie der Naturgesetze mit denen der Freiheit denken sollen, hat etwas an sich, in Ansehung dessen uns eine Wahl zukommt, weil theoretische Vernunft hierüber nichts mit apodiktischer Gewißheit entscheidet, und in Ansehung dieser kann es ein moralisches Interesse geben, das den Ausschlag giebt.

Practical reason selects a better way to think about the possibility of the highest good. It finally postulates God's existence and the immortality of our souls.

## 6. The Possibility of Rational and Mutual Understanding between Cultural Others is Open

The Kantian view of human personhood needs God's existence and the immortality of our souls as the means of promoting morality. But ordinary Japanese people would think they can be adequately moral without such postulations. "Kant in Asia" would also be adequately moral without the postulations. He would not have the same sort of culturally nourished feeling of right and left hands either, insofar as they are originally oriented to Christianity. He would feel something more about natural values.

We distinguish in the moral law the worldview oriented by the principle of cultural value embodiment from the universalizability principle. Only when the universalizability principle is applied to some specific world-description can our view of human personhood and humanity itself be fully completed. This is the concrete embodiment of formal humanity as a self-purpose in concrete personhood. The concrete and materially moral human personhood emerges relatively to moral principles that come from the application of the universalizability principle to the specific world-description. If we see this structure, the possibility of rational and mutual understanding between cultural Others is open. The situation is precisely the same as Kant being able to understand the generous type of Spinozism. Kant selects Christian values. Spinozism selects another view of happiness and morality. They have different kinds of cultural value embodiment and different personalities. These are only contingent facts for the universal pure reason that is common to them

both. But this sort of contingency is not only inevitable but also necessary and of rather great importance.

In his discussion of enlightenment Kant distinguishes “the public use of reason” from “the private use of reason” (8:37). The use of reason by a person functioning as an official of the nation is only the “private use of reason”. He is thinking especially about the officials dealing with religious matters. Following this line of thinking the use of reason adopted for the purpose of creating specific positive culture values would be “the private use of reason”. It involves the discussion of cultural value as a mere fact only, not as the principle of contingent cultural value embodiment. The use of reason adopted for the purpose of mutual understanding of cultural Others, by contrast, would be nothing but “the public use of reason”. We should regard Kant’s categorical imperative as the rational moral principle that enables mutual understanding and cooperation between rationally moral and various culturally different views of human personhood in the one and only, but very rich, world.

## 57. The Idea of Moral Autonomy in Kant's Ethics and its Rejection in Islamic Literature

*Mohsen Javadi*

### 1. Introduction

Kantian ethics alongside utilitarianism and virtue ethics are the three most influential attitudes in modern Western ethics.<sup>1</sup> The history of modern ethics in the West shows us a hot debate between these three alternatives. By contrast, the hot debate in Iran was and still is between those who take morality as a branch of philosophy and therefore a form of rational enterprise and those who take it as a matter of divinity and therefore a form of religious belief. In other words, although the parties of the debate in the West all agree in rejecting or at least neglecting divine command ethics, they disagree about their conception of the source of morality. In Iran, one side of the debate was and still is the proponents of divine command ethics. These scholars did not pay attention to Kant's ethics, not because of some special errors in it, but because of their complete rejection of rationalism. Albeit they also approve of some of Kant's basic points in ethics, like his rejection of relativism, subjectivism, and skepticism or his attention to the role of intention in morality. We, here, put aside the proponents of divine command ethics and concentrate on those who accept morality as a matter of reason.

Utilitarianism from the beginning was neglected or explicitly rejected by Muslim scholars because it treats morality as an experimental science that can be articulated only on perceiving contingent truth in the world. Muslim moralists held and still hold that moral truths are necessary and eternal and hence cannot be grasped by scientific inductive method. It must be noticed that the rejection of utilitarianism as an ethical theory does not imply the rejection of the importance of common

---

1 For a good introduction on these three rival theories, see Marcia W. Baron, Philip Pettit and Michael Slote, *Three Methods of Ethics: A Debate* (Chichester: Wiley-Blackwell, 1997).

utility in ethics. Islamic teachings are full of emphasizing the importance of benefiting others as a moral behavior.

What is currently common among Iranian scholars is a form of virtue ethics that does not contradict the claims of religion. According to this theory, rooted in Aristotle's ethics, virtuous dispositions have an intrinsic value that makes them admirable, but at the same time they are desired for the sake of eudemonia. This character of virtue ethics makes it different from consequentialism, as the latter pays attention only to the end and not to the means. Iranian philosophers have adapted the Aristotelian concept of eudemonia so it includes eternal felicity. This change of the conception of happiness has an important implication for the Muslim understanding of virtues and vices. This is why the list of virtues and vices in the works of Iranian scholars differs from the Greek philosophers' lists and includes many virtues that Aquinas called theological virtues, like faith. Here I do not want to discuss the positive theory of Iranian philosophers but will rather show the negative side of their thinking: their rejection of Kant's ethics. As we know, Kant discussed so many ethical issues that we cannot even outline them in an essay like this. I will restrict myself to one issue that I, following many of Kant's commentators, take as the most important issue of Kant's ethics: moral autonomy. We shall study and cite only those criticisms by Iranian scholars that relate to this issue. I will first define and explain the Kantian thesis of moral autonomy, then talk about its criticisms in contemporary Iran.

## 2. Moral Autonomy

The idea of autonomy can be defined and analyzed in two different ways, both evidenced in the works of Kant. I am aware that there were and still are many controversies about the true and proper interpretation of Kant's idea of autonomy but here I assume the authenticity of the conception of Iranian scholars. I think they were right in this understanding, but I shall not go into a deep discussion of the issue here.

1. Moral autonomy, on one hand, is a theory about the truth conditions of moral judgments; according this theory, moral judgements cannot be determined by any objective truth and principle. This theory, known as self-legislation, is usually classified as a form of constructivist theory in ethics. Although Kant explicitly says that aesthetic judgments about nature are autonomous in this sense, and thus cannot be deter-



mined by any objective truth, we cannot find such explicit statements with regard to moral judgments. Some of his early commentators, like Friedrich Schiller (1759–1805) and Emerson (1805–1882), influenced especially by his idea of self-legislation, held that moral judgments are autonomous in the same sense and thus cannot be determined by objective truth. This interpretation has many defenders in the modern world and continues to be alive in the works of philosophers like John Rawls and R. M. Hare.

2. Moral autonomy, on the other hand, is principally about the conditions of moral agency. In this usage it is typically referred to as the autonomy of moral agents. According to this theory, moral agency is a procedure that begins from the reason-based respect for law and results in the act whose characteristic is to be universal. That is, the person wills it to be acted by all people as a universal law. The entrance of any other external motive, like God, or even internal motive other than reason, like appetites, in this procedure diminishes its moral status and renders it a form of teleological decision making. Respect for moral law, the starting point of moral agency, is rooted in reason alone and cannot be found among the desires and appetites of a person even when they are directed in the same way as moral law demands.

### 3. The Encounter of Iranian Scholars with Kant's Ethics

The history of the encounter of Iranian scholars with Kant's philosophy is very complicated. The first, introductory stage of acquaintance dates back to about one hundred and fifty years ago. At that time one of the Iranian scholars who studied in France asked one of the greatest classical philosophers in Iran, Agha Ali Moddars (1882–1889), some questions about Kant's philosophy. The second stage of the acquaintance begins from 1937 and continues to 1982. During this time some general studies of Kant's philosophy were carried out and Mohammad Ali Foroughi (1877–1943) accurately translated selected works of Kant. But the serious engagement with Kant in Iran is new and relates to the last three decades, when nearly all of his books have been translated into Farsi. Now we can see an increasing number of young scholars who can read Kantian literature in Latin language like English and German. It must be noticed that the other source of Iranian acquaintance with Kant's philosophy in general and his ethics in particular were the

books of some Egyptian writers, because most Iranian scholars (especially those who studied in the seminaries) knew the Arabic language.

Abdel-alla Darraz, who was among the first Egyptian scholars sent to Paris to complete his studies, wrote his dissertation on the code of ethics in the Noble Quran. In his book, *Dustur al-Akhlaq fi al-Qur'an (The Constitution of Ethics in the Qur'an)*,<sup>2</sup> Darraz uses Kant's ideas of ethics with a tone of admiration and approval. Kant's book are cited more than one hundred times and compared with Islamic ethics. On the contrary, another Egyptian philosopher, Youssef Karam (1886–1959), who was under the influence of his famous teacher in Paris, Jacques Maritain, rejected Kant's ethics in the light of his Thomistic background. He argues that Kant, despite his intention to ground morality on a firm base could not succeed in this regard. Abdel-Rahman Badwi (1917–2001) was the last important philosopher who wrote a book on Kant's ethics and explains his ideas in detail and with accuracy. His book is mainly explanatory and contains no comparative or critical points.

#### 4. Aspects of Kant's Theory Rejected by Iranian Scholars

*A. The apriority of moral knowledge.* The first side of the idea of autonomy shows that our knowledge of moral codes is not rooted in any other knowledge we can have about the world. In other words, moral codes are not *a posteriori*, taken from our observation of the world. Moral knowledge is a form of *a priori* judgment that is rooted in the structure of reason alone. Misbah Yazdi (1934-) and Mohammad Reza Modaresi (1955-) criticize this idea in their respective books<sup>3</sup> and argue that all deliverances of practical reason, including moral obligation, can only be based on the theoretical knowledge we have of the world.

Moral language has three partners: the agent, the act, and the end. According to Misbah Yazdi, moral obligation does not refer directly to a relation supposed to be between the agent and the work. Rather, it re-

---

2 Abdel-alla Darraz, *Dustur al-Akhlaq fi al-Qur'an (The Constitution of Ethics in the Qur'an)* (Beirut: Dar al-Kitab al-Islami, 1387/1967).

3 For example, see Misbah Yazdi, MT, *Durus-e Falsafeh-ye Akhlaq (Lectures in Moral Philosophy)* (Tehran: Eteleat, 1374/1995), 91; and Reza Modaresi, *Falsafeh Akhlaq (Ethics)* (Tehran: Soroush, 1371/1992), 219.

fers to the necessary relation that exists between the obliged action and the perfect end of humanity. Indeed, moral necessity is like other necessities in the world: it exists outside of humanity and refers to a relative existential state of affairs. According to Muslim philosophers, we can find this form of relative between each cause and its effect (e.g., between fire and burning). So what makes a relative necessity a moral one is that it refers to the perfect end of human beings. Although this idea was criticized by other Iranian scholars, here we only refer to its point against Kant's idea of autonomy. According to this theory, moral obligation, and so also goodness and badness, all refer to the relations that exist between human perfection as the end and the actions in question. The point of criticism is that morality is not independent of our theoretical understanding about the world and about the causal forces at work there.

*B. The categoricity of the moral imperative.* Youssef Karam was the first Muslim scholar who proposed this criticism, in his book on the history of modern philosophy.<sup>4</sup> After him, some young Iranian scholars developed the idea and criticized Kant in this way.<sup>5</sup> As we know, Kant distinguished between categorical and hypothetical imperatives. Because of its autonomy, morality is based on a categorical imperative that asks us to exercise good will all the way and forever. The point is that our will has two sources and indeed reason asks the will to do what it says if and only if the will chooses to be rational. In other words, reason orders a conditional imperative to the effect that if will wants to keep rational dignity of the soul it must do such and such.

*C. The idea of self-legislation.* It is evident that the *ought* of moral sentences yields the form of legal sentences; indeed, they oblige us to do something or to avoid something. So what is the meaning of self-legislation? Is it possible for a person to oblige oneself? We are familiar with Robert Adams' famous modified divine command ethics, claiming that while good and bad are real things and can be discovered by reason, moral obligation is something that needs a person who has sovereignty over us, like God. This idea is the basis for a critique of Kant's autonomy. This objection is based on a constructivist interpretation of Kant,

---

4 Youssef Karam, *Tarikh al-falsafah al-hadithah (A History of Modern Philosophy)* (Beirut: Dar ol-maaref, 1995), 257.

5 Mohammad Rezaei, *Tabyyen va Naqd Falsafeh Aghlaq Kant (An Explanation and Criticism of Kant's Ethics)* (Qom: Boostan Ketab, 1379/2000).

because in a realist interpretation reason is not the author of the law but the discoverer of it in the real world.

*D. The idea of the formalism in Kant's ethics.* One of the attractive characteristics of Kant's ethics is his extreme formalism, closely related to the idea of autonomy. Karam and some Iranian critics remarked that Kant did not give us the content of moral standards. They hold that his universalism is not sufficient for articulating moral codes. According to this criticism, in many cases we have universal ideas but they do not belong to the moral sphere; also, in some case we have something like supererogatory actions that are not universal but certainly belong to the realm of ethics.

*E. Reason as a unique motivational source.* The other point raised against Kant's autonomy questions why we must make such a distinction between reason and desires and then define humanity in terms of reason alone. It is true that our desires are sometimes in contrary to reason, but this does not imply that we do not have the right to use them in moral motivation at all. Muslim philosophers, usually following Aristotle, hold that for motivation we need desire in addition to some related beliefs. In this line of thought they not only reject the idea of the sufficiency of reason for moral motivation but also emphasize the inevitable role of desires in moral motivation. Kant may say that reason firstly will afford us with the needed desires and then motivate us in the direct way. But contrary to this view some Iranian scholars hold that reason cannot give us desire because it is like light that only shows us things, and our nature is the root of desires that are necessary for working in the world.<sup>6</sup> This is why they try to relate morality to our needs and desires.

*F. The externality of God to human beings.* According to Kant, the derivation of moral maxims from religious revelation and religious commands is contrary to moral autonomy. But the question for Iranian readers of Kant's ethics is why we take reason, which is created by God, to be internal to humanity without worrying about its deliverances, while at the same time taking divine revelation (revealed by the same God who created reason in us) as external to us? According to Islamic teachings, God, although a person and not imminent in us (except in Sufi interpretations), is at the same time internal to human beings: in the words of the Qur'an, "nearer than his jugular vein" (50:16). Respect for rea-

---

6 M.H. Isfahani, *Nihaya al-derayah (Ultimate Understanding)*, vol.2 (Qom: Aal al-Bayt, 1414/1993), 126.

son is a sign of respect for God and we can respect God through obedience to His revelation without compromising the autonomy of reason.

### 5. Concluding Remark

The philosophy of Kant is no longer merely a part of European philosophy, for he has become a conversation partner with philosophers from other traditions as well, including the tradition of Islamic philosophy. Muslim philosophers will benefit from exploring some of the newer interpretations that have been given of Kant's religious views; at the same time, Western philosophers might benefit from conversation with Muslim philosophers who see religion and reason to be so well integrated that the sort of autonomy that Kant sought for morality might be made compatible with a respect and obedience to revelation that might have surprised him.

PART XII

Kant on Asia and  
Asia in Kant



## 58. The Kantian Model: Confucianism and the Modern Divide

*Alain-Marc Rieu*

### 1. Transcultural Kant

In “The crisis of our time” Leo Strauss remarked how difficult it is today to understand Aristotle’s political philosophy.<sup>1</sup> One needs to understand the doctrine of the author in detail, but also how this doctrine is understood and commented on according to the various historical and local conditions of its reception. The result of this translation process is quite distant from the original. This is certainly true of Kant in Europe, and even truer of Kant in Asia. At this level, the challenge is to understand how Kant’s philosophy was reconstructed in Asia in a reception process structured by various cultural traditions and political interests. The problem is not what is left in this process of an original *Ur-Kant* but to define what is this *transcultural* Kant. His philosophy has become the name of a conceptual ideal-type that plays an ideological function, all at once cultural, social, and political.

The theme of this congress, “Kant in Asia”, therefore raises many difficult problems; these highlight the complexities of the present world, of what Edouard Glissant calls “mondiality”, the fact that we all participate in one world and that this world is woven by a diversity of traditions and interests competing with each other. Kant would feel lost in this world because his call for universality, based on theoretical and practical reason, is hard to hear. The problem is not the work of Kant himself but of his multiple receptions. What is received is not the work of Kant himself but interpretations of his work in Western Europe, in those countries where Kant’s philosophy played since the nineteenth century a major role in the formation of various ideologies and philosophical discourses. As a first example, Kant’s work remains

---

<sup>1</sup> In Harold Spaeth (ed.), *The Predicament of Modern Politics* (Detroit: University of Detroit Press, 1964).



the conceptual core of the French conception of the Republic, of the role of Reason in the secularization process, of Human Rights and of the ideal relation of each citizen to public interest and collective morality. France has been a Kantian nation since the late nineteenth century. So studying Kant in Asia is to study the circulation and use of this philosophy both in Asia and the world. In summary, replicating Wittgenstein's famous formula: "don't ask for the meaning of Kant, ask for the use of Kantism today."

## 2. Kant's Function

From the perspective of world history and the history of modern philosophy, Kant is the name of an ambiguous philosophy. Kant has forever associated his philosophy with the Copernican revolution, the formation of modern science and Modernity in general. A proof is the famous comparison between Kant and the French Revolutionary Robespierre, by Heinrich Heine in *Religion and Philosophy in Germany* (1834).<sup>2</sup> In a summary of the *Critique of Pure Reason* Heine explains the opposition between *phenomena* and *noumena* and concludes: "God, according to Kant, is a noumenon. As a result of his argument, this ideal and transcendental being, hitherto called God, is a mere fiction. It has arisen from a natural illusion." For Heine, "this is the sword that slew deism in Germany." As such, Kant is the dividing line between ancient and modern thought, in Europe and also in Asia. He is the philosopher who has limited the pretention of English empirical thought to derive all knowledge from experience, including moral values and even God himself in David Hume's conception of natural religion. From this point of view, Kant is the (German) philosopher who saved morality, religion, and spiritual life from modern rationality in its (English) empiricist version. Kant drew the line between metaphysics and morality on one side and science and modern knowledge on the other side.

But Heine explains also how Kant's critique had an unintended effect in German thought and religion in the nineteenth century:

As in France there were people who maintained that Robespierre was the agent of Pitt, with us (i. e., Germans) there were many who went so far in

---

2 The quotes below are from Heine's *Religion and Philosophy in Germany, A Fragment*, tr. John Snodgrass (London: Trübner & Co, Ludgate Hill, 1882), 107–21.

their willful blindness as to persuade themselves that Kant was in secret alliance with them, and that he had destroyed all philosophic proofs of the existence of God merely in order to convince the world that man can never arrive at knowledge of God by the help of reason, and must therefore hold to revealed religion.<sup>3</sup>

According to this interpretation, Kant's Critical philosophy was paradoxically establishing that metaphysics, morality, and religion were beyond knowledge, as pure objects of faith and creed. Systems of belief could indeed be imagined and constructed beyond demonstration and proof. The *noumena* could not be known in the modern sense because they are not objects given by our senses. But they could exist as pure objects of thought and they could even be reached beyond sensible intuition, by intellectual intuition and pure reasoning. Such philosophies might have no epistemic value, but according to Kant they could have a practical meaning and moral value.

This interpretation of Kant proves the depth and complexity of his philosophy. But in a further paradox, it reinforces the meaning of Kantism for modernity. Saving morality and religion meant and still means saving various types of metaphysics, worldviews, and philosophies. Since Kant, Heidegger has explained to us all that "worldviews" are based on metaphysics, the source of different philosophies.<sup>4</sup> This reading of Kant does not negate or limit a critique of pure reason and pure reasoning, of all Western and non-Western "logocentrism". It implies also that all approaches to metaphysics face the challenge of a Copernican revolution, of evaluating their historical set of beliefs and values. This revolution cannot be reduced to modern science versus pre-modern metaphysics. It concerns modernity's defining criteria: its conception and practice of knowledge. This challenge concerns all cultures that developed within their own local tradition but are now confronted with the question of Modernity. For these cultures, nations or societies, the Copernican challenge is the question of modernity and Kant is the core of this challenge. In this perspective, Kant is the philosopher who drew the limit between modern science and rationality on one side and traditional, national, and cultural identities on the other side. Concrete-

---

3 Heine, 121.

4 See Martin Heidegger, *Being and Time*, tr. J. Macquarrie and E. Robinson (London: Blackwell, 1962), ch. 6, "The Task of Destroying the History of Ontology", 49. See also "L'époque des conceptions du monde" in *Chemins qui ne mènent nulle part (Holzwege)*, tr. W. Brokmeier (Paris: Gallimard, 1962), 119–23; tr. M. Grene as *The Age of the World View*, in *Measure 2* (1951), 269–84.

ly, the result is that one can be morally Confucian and epistemologically modern. This divide and the relation between the two sides of this divide is the true Kantian meaning of modernity. In this sense and until today, Kant saves all pre-modern sets of values and reasoning, worldviews and philosophies. This explains why Kant plays a pivotal role in all discourses on the meaning of modernity.

Kant's philosophy clearly demarcates what comes from experience and can properly be called "knowledge" in the modern sense of science, from another realm of rational discourse situated beyond science. This other realm is what is called "metaphysics". Metaphysics is composed of various beliefs, arguments, historical experiences, values, and discourses that have organized the life of societies and intellectual debates for centuries. In this sense, deeper than a common worldview proper to a given society, metaphysics is the structure of collective representations and everyday life: it determines how individuals and groups understand each other, how they form a society, how this society relates to nature. In this sense, metaphysics is like mythology.<sup>5</sup> It is always located in a given time-space, proper to a given group or society. It defines a nation or a people and what distinguishes one nation from another. Therefore, approaches to metaphysics are many and none of them can pretend to universality or to reach a truth similar to modern science. The philosophy of Kant establishes the ground and value of one such metaphysics within these boundaries: it organizes the common life of a people; it is constructed by reasoning, by a rational use of language that associates ideas, perceptions, and values in well-formed arguments having a practical value and even a collective efficiency. The "world" people are reasoning about within a given metaphysics is itself a product of their collective ways of reasoning. These rational discourses cannot pretend to know the world "as it is". Such a common world is simply a "life form",<sup>6</sup> a form of communication between a certain group of individuals, a collective identity constituting a nation or a culture. But this metaphysics does not satisfy the criteria of truth introduced by modern science.

The "world as it is", outside collective representations organized in various metaphysics, cannot be known "as it is". Still, the outside world

---

5 This conception of metaphysics is derived from Heidegger and also from Derrida. It is also related to the work of J. Lacan, C. Lévi-Strauss, and M. Foucault.

6 In the sense of Ludwig Wittgenstein, for instance in his *Philosophical investigations* (Oxford: Blackwell, 1968), remarks 23, 241.

produces “impressions” on human senses and these “impressions” or “representations” (the “given”) are ordered from the outside. According to Kant, the mind is not a *tabula rasa*: the order of impressions is pre-organized by the structure of mind. But still all knowledge comes from the senses. This is the modern conception and status of “experience”. This is the novelty introduced by modern science. But this novelty cannot be reduced to science. In this sense, the modern experience and modern science irrupted within a metaphysics that was conceived until then as the order of all things, including a conception of nature, God, political order and morality. The modern experience introduced a different cognitive attitude and a different relation to the world: modern humanity was in search of the “real”, simply meaning what is outside metaphysics. By establishing the ground and justification for all metaphysical construction, Kant was establishing also that an exercise of reason beyond experience could not pretend to “truth”. This was a decisive philosophical move. In the Transcendental Dialectic of the *Critique of Pure Reason* Kant defines the structure of the metaphysics that was for him universal, the only one or the only valid one. Beyond experience, reason was organizing all discourses and beliefs under three Ideas of Pure Reason: the self, the world and God. In doing so, Kant was grounding metaphysics and also reducing it to a pure exercise of reason. But because of the reasoning that was grounding it, this metaphysics could not pretend to be universal or true. It was an artifact of pure reason, a projection beyond experience of the structure of the human mind. Metaphysics was defined as the formal structure of a set of beliefs and values. In Kant’s philosophy the content of metaphysics, the shape of these ideals, arguments, and ideas, were proper to European Christianity. But other contents are possible and do exist. To establish such rational ground is to draw the limits of this metaphysics and of all metaphysics. This argument was and still is a challenge for all metaphysics. Is it possible to structure other metaphysics according to the same method, to find both their ground rules and the transcendental ideas, to transform an open an informal set of values, discourses, and creeds into a metaphysics?

The task initiated by Kant was a new sort of philosophy, quite different from all traditions, from the philosophies of Leibniz and his followers, but similar in many ways to the genealogical method<sup>7</sup> introduced by Locke and Hume. This task was itself a product of a “meta-

---

7 This method intends to reproduce by reasoning the imagined formation or generation of concepts or ideas by aggregation of elements.

physical” revolution that gave birth to modern science.<sup>8</sup> Before Kant, metaphysics could be reduced to an organized set of arguments and ideas. It was an endless comment and interpretation of these ideas, but it could not be imagined that all these variations on the same basic ideas could be structured once and for all by reaching the argumentative structure of their presuppositions. Certain variations were producing or assimilating new ideas, but they were always reproducing similar patterns: the ideas of a self, of a world and of a God. Kant’s achievement is to define and expose the structural matrix of all beliefs and discourses in a given society.

Only the cognitive attitude of modern science and its rational requirement can explain the conception and construction of the Transcendental Dialectic defining the patterns of metaphysics. Even religion was reconstructed and redefined “within the limits of *bare* reason”. This explains why Kant is the key modern philosopher. He achieved much more than the separation between experience and metaphysics. He introduced a cognitive attitude and requirement typical of modern science within the construction of rational discourses beyond experience. By doing so, Kant has saturated Western, Christian, and European metaphysics. Until today, his work has been asking us the question: what to do next? What is philosophy beyond metaphysics, in an age characterized by the emergence of science? What interested Kant was not the sciences themselves but a new cognitive practice associating experience and reason, the two sources of human intelligence: the given from sense and the given from mind. From this point of view, “Kant in China” is a challenge to all metaphysical constructions and philosophies.

Historically, Kant is a lock as well as an open door. He was separating modern science and metaphysics. But he was also unlocking the role of science as an autonomous field of rational activity, of interaction between experience and reason that had long-term consequences for all human intellectual activities. By doing so, Kant was reducing the sphere of religion and metaphysics to the sphere of morality. He was providing a ground for an autonomous sphere of morality. Morality became the substitute of religion and metaphysics, the truth of mankind beyond knowledge and science. Kant was ending a conception of religion as the final discourse on the world and history, expressing a transcendent

---

8 “The metaphysical foundations of modern Science” have been studied by many historians of science and philosophers. This is also the title of a famous book by E. A. Burtt (New York: Doubleday, 1924).

principle dictating the ends and proper means of humanity. Beyond science, all was reduced to morality. His conception of history and humanity was an exercise in reason, experience, and morality. This was for him the meaning of the Copernican revolution. It was a step without return in the evolution of humanity. For Kant, it was a scientific and philosophical revolution, a revolution for humanity and world history. Kant was locking in the old world, the world of traditional metaphysics, and opening a new world.

This explains why Kant is now the cornerstone of modern philosophy. But Kant's revolution also had its counter-effects. His philosophy opened a new world but it also unlocked a deep nostalgia for the old world, where knowledge and morality, reason and experience, religion and understanding were united in a common metaphysics. Many of Kant's successors conceived their work as a stitch closing this gap, or wound, between the metaphysics of pre-modern Europe and the world of reason, science, and experience. Continuity had to be found and proven. This was how Hegel imagined his historical role and duty as a philosopher. In Europe emerged various ideologies and philosophies that were denying and even rejecting the modern world from the point of view of politics and morality. The world had lost the ground needed to assure a moral order, an ordered society, where each individual could find a place and duty in a perspective extending beyond the short view of private interest. Science and modern culture were conceived as a recent (modern) mistake, as a narrow and shallow conception of the world and humanity.<sup>9</sup> This is Kant's historical paradox: his philosophy unlocked a new world; but this world closed the pre-modern world, thus refusing this new status and, in the name of morality, rejecting religion and metaphysics, the world of science, individuality, and experiment. The world could not be an endless experiment. It had to be kept within a common historical worldview. Nothing new could or should happen that was not a re-interpretation of the same, of the old and the past.

---

9 See, for example, the work of Charles Baudelaire, in particular his *Écrits esthétiques*, and Walter Benjamin's comments on Baudelaire.

### 3. The Kantian Effect

There are no laws of history for social and cultural processes, just observable regularities. But some philosophical or cultural ideal types circulate, are adopted and adapted. Because of its historical weight, Kant's philosophy seems to express an evolution that has been observed in European societies and can therefore be expected in others. Kant's philosophy has played a major role in contemporary Chinese philosophy and culture. Typically, according to Lee Ming-huei, "Mou Tsung-san's interpretation of Confucianism is characterized by the influence of Western philosophy, especially that of Kant. In Mou's interpretation, he employs (...) his philosophical framework of 'appearance' and 'thing-in-itself'."<sup>10</sup> Fabian Heubel explains why "in the preface of his book *Intellectual intuition and Chinese philosophy* (1971), Mou Zongsan bluntly states that the whole of Chinese philosophy would have been in vain if the possibility of intellectual intuition could not be proved."<sup>11</sup> For Mou Zongsan, Kant's thought is the ideal type of modern philosophy because of its radical opposition between practical and theoretical reason, between science considered as modern universal knowledge and morality considered as beyond knowledge. More generally, Kant's practical philosophy demonstrated the full validity of moral judgment as well as the universality of moral law beyond all content. Moral law is considered as a form grounding morality beyond any empirical values or behaviors. This divide between theory and practice, being the frame of the Kantian model, is the source of its influence in East Asia and elsewhere. It is the criteria and operator of the modern.

Interestingly, Hegel's philosophy cannot perform this function because his thought pattern intends to overcome the opposition between theory and practice, knowledge and morality. In the case of Chinese culture, the Hegelian ideal-type implies an historical modernization

---

10 Ming-huei Lee, "Mou Tsung-san's Interpretation of Confucianism: Some Hermeneutical Reflections" in Ching-I Tu (ed.), *Classics and Interpretations: The Hermeneutic Traditions in Chinese Culture* (New Brunswick and London: Transaction Publishers, 2000), 411–25.

11 Fabien Heubel, "Transcultural Explorations into Contemporary Philosophy of Cultivation: On the Critique of Kant in Mou Zongsan and Foucault" (unpublished manuscript), part 2. See also the introduction by Joël Thoraval to Zongsan Mou, *Spécificités de la philosophie chinoise* (Paris: Le Cerf, 2003), and Anne Cheng, "Les tribulations de la 'pensée chinoise' en Chine", in Anne Cheng (ed.), *La pensée chinoise aujourd'hui* (Paris : Gallimard, col. Folio, 2007).

project that leads to overcoming Confucianism as the defining character of Chinese morality and identity. Such a new or *modern* cultural identity has been at stake in China since the end of the nineteenth century and until today. But this is a false and destructive conception of the modern divide. The Kantian model goes deeper into the Chinese modernity dilemma because the modern divide is asserted and not repressed. This divide is the condition for modernizing China while preserving Confucian values. The Hegelian model is not *transcultural* but mono-cultural: it develops a cultural monism.<sup>12</sup> Marxism is a version of the Hegelian conceptual ideal-type. Today, the shift from a Hegelian ideal type to Kant shows that the modern divide cannot be or should not be overcome.

Moral law, because it was the form and the ground of all morality, can be discovered, established, and obeyed according to various historical circumstances or cultural conditions. The real sense of moral law is not its pure or abstract form but the shape this form can take according to life's diversity. In a transcultural Kantian perspective, this shape has to be found through a personal process enabling an individual to develop a relation to her or himself at a distance from empirical daily life by experiencing a morality dilemma. At this point, the problem of morality transforms into the problem of "self-cultivation", of ethics understood as the relation of one's self. This means that Confucian practices might today retreat into tradition but in another way they have a significant role to play in a modern society, in the modernization process and in a modernized society. The transcultural meaning of a Kantian conception of morality cannot be reduced to moral law: it is the process whereby morality is conceived, practiced, and established. The goal of this search is certainly to overcome all empirical circumstances but, beyond the law itself, the actual shape of the moral law and effective meaning, the way self-cultivation is achieved and performed, differ according to each society and circumstances. In China, the problem is to imagine a new conception and practice of morality, in order to save or protect Confucianism from the destructive aspects of Chinese history and also to invent and teach a modern or new Confucianism. This is the explicit task of many contemporary Chinese philosophers. It is essential for Chinese society and culture. Due to the Kantian ideal type,

---

12 Hegel's cultural monism, his conception of being, logic, and history, finds its root explicitly in Christianity, as expressed at the beginning of the Gospel of Saint John.



it is possible to reconcile modern universal science with Sino-Confucian morality considered as the historical ground and present meaning of Chinese culture. Finally, Confucian morality becomes also the norm for evaluating individual and collective behaviors. This potentially also has a strong political meaning.

Finally, the modern divide established by the Kantian model leads to a new conception humanity. In Kant's philosophy and in a European conception of mankind, the divide between knowledge and morality gives in the end preeminence to knowledge. Knowledge activities are the basis of the Western-European conception of the role and responsibility of humanity. Kant typically constructs a definition of moral law as the form of all morality systems according to a scientific method. The relation of the individual to her or himself is determined by the recognition of the preexistence of a law and by submission to this law, be it moral or natural. Morality tends to be viewed as submission and obedience to law. Individuals live their private life in the constant shadow of moral law and the resulting guilt feeling.<sup>13</sup> But reinterpreting such a Kantian conception of morality according to Confucian ethics opens a variety of subjective practices. Ethics becomes a personal quest enabling given individuals to free themselves from daily life, from painful situations or repressive social conditions in order to achieve a state of "contentment" by themselves and within themselves, whereby they can relate to others or to the world in general with a different attitude, free from rivalry, envy, or revenge. In these conditions, practical morality and self-cultivation have necessarily an impact on knowledge activities. They should have the potential to influence humanity's relation to the world, to nature, in a sort of pragmatism. If the Kantian model submits morality to a scientific attitude, a Confucian model should ideally be able to develop a collective ethics opening a different conception of knowledge, even of science and technology, of doing research and experimenting in general. By contrast, in a Hegelian paradigm, an Asian or Confucian ethics would just be a moment in the effective construction of science and morality at the service of the State. Reinterpreting Confucianism in a Kantian model leads to quite a different perspective.

Such comments derive not from the historical Kant but from a transcultural Kantism. Such reinterpretations at the common border of

---

13 This is historically the reason Nietzsche's critique of the conception of morality opened a new philosophical era in Europe. "Nietzsche in China" might be too early.

a Kantian model and Confucian ethics show the vast ideological constructions being undertaken in the West and the East. But they prove also the wealth of philosophical innovations taking place at this border. These innovations have a deep impact on ideological constructs. What is at stake is the relevance of the so-called Asian values as taking historically the relay of Western values in Asia and potentially in the whole world. The idea of an historical relay and potential overcoming suppose a typical Hegelian model. As already shown, the potential relevance of a dominant Confucian morality (i. e., of an empirical example of a moral system asserting its universal value) can only be grounded in a Kantian model, not in a Hegelian one. Therefore Asian values, as a discourse, cannot pretend to overcome anything. They are a legitimate moral claim as long as they are proven a valid case of morality from a transcultural perspective. In order to be valid, this claim requires recognizing the diversity of value systems, all pretending to express a universal sense of morality. In this case, “universal” simply means “general” and “grounded”. But it does not refer to any transcendental or transcendent ground. If Confucian values are not understood as a set of conventional customs proper to a given local culture and society, they require a Kantian ideal-type in order to establish a proper ground and be recognized as a full value system participating in world morality. Therefore, Confucian values cannot be understood or promoted in an overcoming mode without becoming an object of deconstruction. When adopted and reconstructed, the Kantian model provides an in depth structure for such debates and ideas.

Still, a problem remains. Kant needs to be understood also as the philosopher who performed the final reduction of Christianity to a given moral system. By doing so, Kant’s philosophy was not only fabricating and grounding modern European morality. On a larger scale, this philosophy was reducing the whole Christian historical worldview to a morality. This reduction operated because modern science and its general cognitive attitude had constructed another worldview. According to Kant, the former worldview was Christian and included a conception of society, politics, and mankind in their relation to a totality ordered by the Christian conception of God and the related belief institutions and practices. Obviously, implicitly or not, Kant’s philosophy has mutated Christianity into a new conception of morality, the mutation of the Christian worldview into a new conception of personhood, law, society and history. The distinction between theoretical reason and practical reason, between science and morality, proves that this mutation

was, for Kant, irreversible. From the modern divide there is no return. It could not be “overcome” in the Hegelian sense, so that the old world was saved and preserved within the new world.

#### 4. Conclusion

The historical adoption and adaptation of Kant’s philosophy in Asia imply that what happened to Christianity and to the Christian world is also happening or could happen in the Sino-Confucian world. This is the historical role played by those philosophers in China who are searching in Kant for the key to a modern China, to a modern Chinese world, with the goal of saving Confucianism and an “imagined” Confucian world. This task was and is clearly a rejection of a Hegelian conception of modernization in its Marxist form that led in the past to the program of overcoming Confucianism and building a new China where humanity, economy, and politics were integrated into a “one national community, one society-economy, one party” system. The return to Kant, in order to abandon Hegel and therefore Marxism, is supposed to save Confucianism. But Confucianism is saved by being deeply transformed. To transform or reinterpret Confucianism has as its consequence, or even its goal, transforming Chinese culture and society. This is the real issue raised by this transcultural Kant fabricated in Asia. This return to Kant goes far beyond philosophy and Confucianism. The effective goal is to conceive and build the conceptual ground of a new China based on the divide between Confucian morality and a *modern sphere*, external to morality, exactly as the realm of modern knowledge was for Kant outside the realm of practical reason.

## 59. Asian Hospitality in Kant's Cosmopolitan Law

*Klaus-Gerd Giesen*

### 1. Introduction

Immanuel Kant was never in Asia. He hardly ever left Königsberg, today's Kaliningrad in Russia, where he was born in 1724 and died in 1804. According to his biographers, Kant never departed from the duchy of Prussia.<sup>1</sup> Nevertheless, the philosopher mentions Asia on several occasions in his work, especially in *Perpetual Peace* (1795) and *Metaphysics of Morals* (1797). In the former book he calls attention to the situations in East India (Hindustan), China (including Formosa and Tibet) and Japan,<sup>2</sup> while two years later he points to New Holland, today's Australia (6:353). In all these cases Kant uses Asian examples to illustrate his view on European colonization and to derive the cosmopolitan law literally from them.

One could ask how it is possible that Kant was so well informed about Asia and other continents, while he always stayed in his birthplace and surroundings? In the first place the philosopher read, of course, many academic studies. For instance, in *Perpetual Peace* he quotes Alaph, Fischer, Hesychius, and Horatius as sources of information (8:359). In addition, Kant gathered valuable insights about all regions in the world from the incoming seafarers and traders. Königsberg, a big town of about 60,000 inhabitants (more than Munich and Frankfurt at that time) that had been member of the Hansa League, was one of the main ports of the Baltic Sea; ships came in from England, Holland,

---

1 His furthest travel took him in the fall of 1765 to Goldap, a small town a few miles from the (then) Russian border and about 120 km from Königsberg. From 1748 to 1755, he was a private teacher in Judschen and Groß-Arnsdorf that were less than 100 km distant: Arnulf Zitelman, *Nur daß ich ein Mensch sei. Die Lebensgeschichte des Immanuel Kant* (Weinheim/Basel: Beltz, 1996), 64 and 128–9.

2 See 8:358–9. Kant devotes a long footnote to Tibet, at his time part of Qing China while being twice occupied by Nepal, and depicts it as being historically the main joint of trade between Europe and China.

Scandinavia, Poland, Russia and other places, importing among other things Russian commodities, English handcraft as well as colonial products, and above all fresh news (including foreign newspapers).<sup>3</sup> New perspectives on world affairs may have been brought about during the Russian military occupation of Königsberg that lasted from 1758 to 1763. Kant taught mathematics, architecture and pyrotechnics to the Russian officers, and participated in many mundane events organized by the Russian aristocracy.<sup>4</sup> Another crucial influence on his growing cosmopolitan worldview is mentioned by Karl Vorländer: the daily discussions with his best friend, the English businessman Joseph Green from 1768 on, as well as regular meetings with other foreign businessmen residing in Königsberg (Robert Motherby, George Hay, etc.).

## 2. Colonialism Versus Cosmopolitan Law

According to Vorländer, it was likely with Joseph Green that Immanuel Kant had his heaviest disputes about European colonialism, especially at the time of the English–American war of 1776–1783 that led to the independence of the thirteen American colonies from King George III: Kant passionately defended the secession, while the Englishman Green strongly opposed it.<sup>5</sup> It is therefore not surprising that the question of colonialism is at the heart of Kant’s cosmopolitan law *stricto sensu*, elaborated in the third Definitive Article of *Perpetual Peace* and §62 of *Metaphysics of Morals*. While the first Definitive Article of *Perpetual Peace* and §51 of the *Doctrine of Law* deal with the internal, constitutional form of the state, claiming it should be republican in order to make it as peaceful as possible (the axiom of so-called “Democratic Peace”), the second Definitive Article of *Perpetual Peace* and §§53–61 of the *Doctrine of Law* aim to lay out the normative foundations of the relations between states, including just war principles and the ultimate *telos* of a worldwide *foedus pacificum*.<sup>6</sup> Significantly, the question of colonialism is dealt with in the third Definitive Article of *Perpetual Peace* and §62 of the *Doctrine*

---

3 Ludwig von Baczko, *Versuch einer Geschichte und Beschreibung Königsbergs* (Königsberg: Goebbel und Unzer, 1804), 96–7.

4 Zitelman, 98–100.

5 Karl Vorländer, *Kants Leben* (Hamburg: Felix Meiner Verlag, 1911), 61–2.

6 Cf. Klaus-Gerd Giesen, *L'éthique des relations internationales* (Bruxelles: Bruylant, 1992), 166–7.

of *Law*, both devoted to the cosmopolitan law (i. e., the relations between domestic individuals and foreign peoples [*Völker*] and foreign individuals). This is the part of his philosophy of international relations where Kant uses Asian case studies, so to speak, to make his point.

It should be underlined that the very base of the cosmopolitan law *stricto sensu* is, thus, the regulation of connections between individuals and peoples, not between states. The reason seems rather obvious: cosmopolitan law *stricto sensu* is all about questionable colonialism; and since, by definition, colonies are not independent states, Kant employs the concept of people. We should, indeed, not forget that the number of states in the international system of the second half of the eighteenth century was rather limited, as most parts of the world were either not yet “discovered” by the Europeans (for instance inner Africa) or else colonized by them. The validity of the second Definitive Article is hence confined mainly to the relations among European states. Kant understood that the colonial question could not be handled by setting it within the classical interstate framework. Consequently, the *ius cosmopolitanum* becomes a construction based on individuals and peoples.

Kant sees the relations between sovereign states primarily embedded in the long history of the European continent, while the non-European sphere of the world raises the ethical and political question of colonialism. For the former he develops a complete theory of *foedus pacificum* in *Perpetual Peace* and a theory of just war in *Metaphysics of Morals*: the republican form of government may eventually spread to the entirety of Europe and North America, and later hopefully to the rest of the world. Until it reaches a worldwide extension war will remain, either between non-republican states outside the *foedus pacificum* or between one or several republican states and one or several non-republican countries. For this reason a theory of war is absolutely required as a complement to his peace teleology. Kant adopts the tradition of just war theory to assess ethically the justice of each war. However, he introduces two major modifications: on the one hand he goes back to the just war criteria defined by Aquinas and ignores especially the *jus in bello* criterium introduced by Vitoria (6:346–8; §§56–7). As I have demonstrated elsewhere, Kant believed the rise of mass war with the Napoleonian wars and the invention of the *citoyen-soldat* and of *mobilisation générale* made it henceforth impossible to distinguish between combatants and civil population.<sup>7</sup> On the other hand, Kant adds to *jus ad bellum* and *jus in*

7 Klaus-Gerd Giesen, “Kant et la guerre de masse”, in Société hellénique d'études

*bello* a new *jus post bellum*, stipulating that a war can be labeled as just only if it is declared and fought by taking already into account the situation that has to prevail once victory is achieved.<sup>8</sup>

As already emphasized, these theories of progressive pacification and of just war, theories I shall not detail here, apply exclusively to the relations between states. The interactions between states and their respective colonies call for a different approach. Here Kant chooses a historical argumentation: his starting point is the arrival of one single individual on the territory of a foreign people (*Volk*). According to Kant, a people is, ethically speaking, always *already* sovereign, even if, from the point of view of positive law and power politics, the state has not (yet) been created. He declares: “Nature wisely separates peoples, which the will of any state ... would be to unite by ruse and violence ...” (8:368). An additional sign that the ultimate (ethical) sovereignty is embedded in the people, not in the state, lies in the fact that the issue is tackled in the last Definitive Article of *Perpetual Peace* and the first paragraph of the *Doctrine of Law*, both explicitly presented as the final achievements of the two works. State sovereignty merely derives from the original sovereignty of the people.<sup>9</sup> Such an interpretation of Kant’s thought avoids, thus, today’s common divide between globalists (Beitz, Pogge, etc.) and communitarians (Walzer, Taylor, etc.) by introducing the concept of people as the key to cosmopolitan law, rather than individuals or the state.

The earth being a “*globus terraqueus*” that territorially cannot be extended, no people has *a priori* more rights than any other to live on a part of the planet’s surface. Since there is a “*Gemeinschaft des Bodens*” (community of the [earth’s] land) that according to Kant is not a community of possession (*communio*), each people must respect the others’ sover-

---

philosophiques (dir.), *Droit et vertu chez Kant. Actes du IIIe Congrès de la Société internationale d’études kantienne de langue française* (Athènes: Union scientifique franco-hellénique, 1997), 334.

8 Klaus-Gerd Giesen, “L’actualité de la philosophie kantienne de la guerre”, *Arès*, XVII.43 (juillet 1999), 73–4.

9 In the classification of Peter Burg this concept of sovereignty corresponds to the natural law definition à la Rousseau: Peter Burg, *Kant und die französische Revolution* (Berlin: Duncker & Humblot, 1974), 165–84. See also: Ricardo Terra, “La souveraineté populaire chez Kant: de l’idéalisme au procéduralisme”, in *Société hellénique d’études philosophiques* (dir.) 78–81. It seems that John Rawls took over this conception in his seminal *The Law of Peoples* (Cambridge, Mass.: Harvard University Press, 1999), v (Preface).

eignty over the land (6:352 [§62]). Colonization beginning historically with the “discovery” of new territories by European peoples—for instance, New Holland (today's Australia) “discovered” by Thomas Cook during Kant's lifetime—the first step is to clarify whether such a first contact initiated by a few individuals of one (generally European) people is morally acceptable. The problem is certainly not innocuous, as demonstrated nowadays by several tribes of the Amazon basin in South America and of the Andaman Islands in the Indian Ocean, which refuse *any* contact with the so-called civilization. Kant strongly believed that it should be permissible to initiate such a first contact: it is “the right of a stranger not to be treated as an enemy when he arrives in the land of another [people]” (8:358).

However, once the contact is established between individuals (as members of different peoples) the situation changes: after having offered sociality (*Gesellschaft*) to the people they are visiting, the “intruders” can be sent away and further contact can be declined, except if this would lead to their “fall” (*Untergang*). Here Kant introduces a special clause for political and religious refugees, probably having in mind the fate of the many Huguenots in East Prussia of his time.<sup>10</sup> Visitors enjoy the cosmopolitan right to stay as long as they are threatened in their home country and as long as they behave peacefully. Otherwise, hospitality, the very base of Kant's cosmopolitan law, does not imply the right to be a permanent guest (*Gastrecht*). In addition to asylum, hospitality is strictly limited to two scenarios: first, the right of an individual not to be treated malevolently upon arrival and to *try* to offer his sociality (an offer that can be refused); second, the right to commerce, as strictly confined to the establishment of intercommunications for trade purposes (for instance, in ports such as Königsberg). This liberal bias was certainly influenced by his best friend, the English trader Joseph Green, but possibly even more by the Scottish moral philosopher Adam Smith, founder of modern economics and father of the liberal metaphor of the “invisible hand”, who published *The Wealth of Nations* just 19 years before Kant wrote his *Perpetual Peace*. Kant claims (8:358): “In this way distant parts of the world can come into peaceable relations with each other, and these are finally publicly established by law. Thus the human race

---

10 Most Protestant inhabitants of Judschen, where Kant was a private teacher from 1748 to 1751, came originally from France and Swiss Romandie. Cf. Zitelman, 65.



can gradually be brought closer and closer to a constitution establishing world citizenship.”

For all the cases that exceed these three fundamental but strictly limited rights to hospitality (asylum, visits, trade), Kant makes it very clear (8:358) that a “special beneficent agreement” (*besonderer wohlthätiger Vertrag*) has to be arranged between the two equal parties, defining the conditions of residency (*ius incolatus*) (6:353, §62). In the absence of such an accord the residency of a foreigner is a moral wrong. It becomes an injustice toward the local people if a *state* organizes the massive *accolatus* of its citizens on the territory of a foreign people. Kant has no mercy when it comes to evaluating the colonizing practices of European states of his time: he emphasizes (8:358) “the inhospitable actions of the civilized and especially of the commercial states of our part of the world. The injustice which they show to lands and peoples they visit (which is equivalent to conquering them) is carried by them to terrifying lengths.” The judgment is severe and limpid: colonization is a violent abuse of the three rights to hospitality, and is therefore a grave injustice. Like several philosophers before him (Las Casas, Vitoria, etc.), Kant clearly takes the side of the oppressed, colonized peoples against the European imperialists.

### 3. Asia in the Cosmopolitan Law

At this point of Kant’s argumentation the empirical situations in Asia come into the picture. In addition to denouncing the “most cruel slavery” taking place in the Caribbean, Kant almost flies into verbal rage when he evokes the many colonial injustices as a consequence of the European “discovery” of “lands without owners, for they counted the inhabitants as nothing” (8:358). In particular, the case of East India (Hindustan) strikes him as a flagrant offense to hospitality granted to Europeans (8:358–9): “under the pretense of establishing economic undertakings, they brought in foreign soldiers and used them to oppress the natives, excited widespread wars among the various states, spread famine, rebellion, perfidy, and the whole litany of evils which afflict mankind.”

Kant seemed very well informed about the fact that after the collapse of the Mogul empire several French and English trading posts were established on the East Indian coast in the early nineteenth century (Pondichéry, Yauaoun, Karikai, and Chandernagore by France, and Madras

and Calcutta by England), either by the *Compagnie française des Indes orientales*, founded by Colbert, or by the British East India Company. They were to become rival colonial territories (in addition to Dutch occupied Ceylon). Both European countries, France and England, were striving for predominance on the Indian subcontinent, and finally heavily clashed in a major war in 1757. With the victory on the battlefield of Plassey near Calcutta, Great Britain gained over France the exclusivity of colonial rule in East India. France kept only the trading post in Pondichéry.<sup>11</sup> The war was entirely imported from Europe, and the Indian population suffered a lot from it, as Kant correctly points out. He was neither on the side of the British nor of the French, but instead indirectly supported the Indian peoples, notably the Maratha Confederacy.

Kant's anger against European colonialism, perceived as being opposed to a liberal world system based on free and peaceful trade among self-determined peoples, is not less present when he examines the situation in China and Japan. He observes that both countries already had some experiences with European colonial powers and were able to draw inferences from these:

China and Japan (Nippon), who have had experience with such guests, have wisely refused them entry, the former permitting their approach to their shores but not their entry, while the latter permit this approach to only one European people, the Dutch, but treat them like prisoners, not allowing them any communication with the inhabitants.<sup>12</sup>

Kant deliberately uses the concept of guest to refer to the European presence in China and Japan. The Portuguese were the first Europeans to arrive in Canton (Guangzhou), establishing a monopoly on the external trade out of its harbor by 1511. In 1535, Portuguese traders were allowed to anchor ships in nearby Macau's harbor and were granted the right to carry out trading activities, though not to stay onshore. Not until 1557 did the Portuguese establish a permanent settlement in Macau. China retained sovereignty and Chinese residents were subject

---

11 P. J. Marshall, *Bengal: The British Bridgehead, Eastern India 1740–1824* (Cambridge: Cambridge University Press, 1988), ch. 3.

12 8:359. Kant also mentions the special case of New Holland (Australia), recently discovered and used as a deportation camp for prisoners expelled from England. He states that even the good intention to purify the home country of bad individuals does not justify the injustice of employing colonial means (8:354 [§62]).

to Chinese law, but the territory was under Portuguese administration. In 1582 a land lease contract was signed, and annual rent was paid.

Despite Chinese efforts to keep European traders and citizens within the area of Macau, European trade spread in the sixteenth century throughout China and threatened virtually to take over the country (just as in India). The so-called “Canton System” finally limited the number of ports where European traders could do business in China. It also prohibited any direct trading between European merchants and Chinese civilians. Instead, the Europeans, generally employees of the major trading companies, had to commerce with an association of Chinese merchants known as the *Cohong*. The European presence was restricted to the harbor of Canton during the trading season, but the foreign traders were permitted to remain on Chinese soil at Macau in the off-season.<sup>13</sup>

Kant seems to have known this situation perfectly and taken it as a model for his normative theory of hospitality, as well as the special case of Dejima (Dutch: Deshima) under the isolationist *sakoku* (“closed country”) policy during the Edo period in Japan. The artificial island Dejima, constructed in 1634 following the orders of shogun Iemitsu, originally accommodated Portuguese merchants. Four years later the Portuguese and other Catholic nations were expelled from Japan, but not the Dutch *Vereenigde Oostindische Compagnie* (*VOC*). However, the shogunate ordered the Dutch to transfer all its trading operations to Dejima in 1641. From then on, only the Chinese and the Dutch could trade with Japan. It is noteworthy that Dejima was an artificial island, and hence not part of Japan proper. Thus, the foreigners were kept at distance from the soil of Japan. Dejima was a very small island, 120 by 75 meters, linked to the mainland by a small, guarded bridge. The trade monopoly with Japan was very profitable for the *VOC*, initially yielding profits of 50% or more. After its bankruptcy in 1795 (the year Kant published *Perpetual Peace*), the Dutch government took over the settlement. For two hundred years, the Dutch merchants were not allowed to enter Japanese territory, and Japanese were likewise banned from entering Dejima.<sup>14</sup>

---

13 Louis Dermigny, *La Chine et l'Occident: le commerce à Canton au XVIII<sup>e</sup> siècle, 1719–1833* (Paris: Sevpen, 1964).

14 Charles R. Boxer, *Jan Compagnie in Japan, 1600–1850* (Oxford: Oxford University Press, 1968).

While for the philosopher of Königsberg Macau and Dejima remained a model to preserve both the Asian peoples' sovereignty and the Portuguese and Dutch right to trade and to offer their sociality to the Chinese and Japanese, Dutch imperialism in Formosa (Taiwan) was the counter-example of a mutually beneficent agreement between the intruders and the local people. The period of Dutch colonial government on Formosa lasted from 1624 to 1662. The *VOC* established its presence on Formosa to trade with China and Japan, and to hinder Portuguese and Spanish trade and colonial activities in East Asia. Indeed, at the beginning of the seventeenth century the forces of Catholic Spain and Portugal were in opposition to those of Protestant Holland and England, often resulting in open warfare in their possessions in Asia. The Dutch first attempted to trade with China in 1601, but were rebuffed by the Chinese authorities, who were already engaged in trade with the Portuguese at Macau. They then built Fort Zeelandia, a fortress on Formosa. On the Southern coast of the Formosa island more and more villages came under Dutch colonial domination, mainly through military action. However, the north of the island was since 1626 under Spanish control. In 1642, the Dutch sent an expedition of soldiers and aboriginal warriors there in ships, managing to drive out the small Spanish contingent from Formosa. From then on the Governor was empowered to legislate, collect taxes, wage war and declare peace on behalf of the *VOC*, and therefore by extension the Dutch state.<sup>15</sup> In 1661, a naval fleet of 1000 Chinese warships attacked Formosa in order to oust the Dutch from Fort Zeelandia. Following a nine-month siege, the Dutch were defeated and the colonial system brought to an end.

Macau and Dejima were the two positive empirical examples, Formosa and Hindustan the two negative ones, that Kant explicitly mentions when devising, in 1795, his cosmopolitan law based on the right to hospitality. It is true, however, that after Kant's death in 1804 the whole situation completely changed, notably with the first opium war and the British occupation of Hong Kong in 1841, transforming the complex relationships between the Europeans and Asians into a nightmare.

---

15 Hsin-hui Chiu, *The Colonial "Civilizing Process" in Dutch Formosa, 1624–1662* (Leiden/Boston: Brill, 2008).

#### 4. Universalism and Interventionism

Presumably, the abuse of Asian hospitality in East India and Formosa by the British, French, and Dutch, contrary to its respect at Macau and Dejima by the Portuguese and the Dutch (as well as by the Jesuits in China in general), led Kant subtly to distinguish between the three rights to hospitality (asylum, visits, trade) and intolerable colonial behavior. The latter must be strongly condemned on the ground of cosmopolitan law. Furthermore, in Kant's view, colonialism is such a strong abuse of cosmopolitan hospitality that he felt obliged to add in the third Definitive Article of *Perpetual Peace*: "... the narrower or wider community of the peoples of the earth has developed so far that *a violation of rights in one place is felt throughout the world*" (8:360, my emphasis). This often quoted sentence about the world community and the violation of rights in another part of the world, very much highlighted nowadays in all discussions about international public opinion and universal human rights, hence stems directly from Kant's empirical analysis of European colonialism. It takes the form of a justification: *because* a world community of peoples exists, a serious violation of rights by one people or state is an attack against all other peoples.

Kant goes even one step further by stating (8:360) that the idea of cosmopolitan law "is a supplement to the unwritten code of the civil and international law, indispensable for the maintenance of the public human rights and hence also of perpetual peace." He makes clear that the colonial abuse of the right to hospitality is considered to be a violation of universal human rights, these being themselves a precondition of perpetual peace. No durable peace can be established without respecting basic human rights. Colonialism, in Asia and elsewhere, is a major obstacle to perpetual peace.

At this point two issues remain. First, is this form of universalism not a naturalistic fallacy? Does Kant not deduce the universal norm from a fact (i. e., from the existence of an international public opinion, or, as Stéphane Chauvier says, a "cosmopolitan consciousness"<sup>16</sup>)? The answer to this question is actually easy to give: Kant certainly distinguishes between, on the one hand the *universalizability* of the human rights norm, based on the categorical imperative, and on the other the fact-driven *universality* of the feeling of its violation. Were the latter not to exist,

---

16 Stéphane Chauvier, *Justice internationale et solidarité* (Nîmes: Editions Jacqueline Chambon, 1999), 141–51.

because of an insufficient worldwide communication system, the human rights violation itself would nevertheless still subsist; it would just not be known elsewhere. Therefore, the very existence of the world community of peoples through means of communication is a mere *additional* and factual precondition for the establishment of perpetual peace. Kant himself characterizes it as a “complement”.<sup>17</sup>

The second issue is about the consequences to be taken by a people called upon by a severe abuse of the right to hospitality elsewhere in the world. Should it militarily intervene, for instance, in order to put an end to a colonial injustice? The answer is given in the fifth Preliminary Article of *Perpetual Peace*: Kant's theory of war stipulates that a *just* war can only be waged as a response to a direct military aggression or threat by a hostile state.<sup>18</sup> What is called nowadays a humanitarian intervention is therefore morally not admissible. However, that does not mean that according to Kant colonialism and other massive violations of basic human rights should merely be observed without taking *any* action. Passivity would be a breach of the moral requirement to seek the “continuous approaching” (*kontinuierliche Annäherung*) of perpetual peace (8:360). It just means such an action should remain non-violent. The idea that the liberal philosopher Kant most likely had in mind is the following: all other peoples and individuals should express their concern and their sympathy for the victims, morally condemn the perpetrating state, and take peaceful counter-actions—for instance, through economic non-cooperation and sanctions, as well as diplomatic pressure and isolation—whenever a severe violation of fundamental human rights occurs elsewhere, such as in ... China.

---

17 On the “complementary cosmopolitanism” see the interpretation by Otfried Höffe, *Wirtschaftsbürger, Staatsbürger, Weltbürger* (München: Beck, 2004), 152–3.

18 8:346. See also Véronique Zanetti, *L'intervention humanitaire* (Genève: Labor et Fides, 2008), 36.

# 60. Doing Good or Right? Kant's Critique on Confucius

*Rein Vos*

## 1. Introduction

Kant and Kantian philosophy have received much inspiring attention from Chinese philosophy.<sup>1</sup> However, Kant and his philosophy have been qualified in opposite directions, positive and negative. Nietzsche named Kant, perhaps cynically, “the Chinaman of Königsberg”<sup>2</sup> and in this line of thinking several authors have acknowledged the similarities in philosophy and daily life between Kant and Confucius.<sup>3</sup>

Amongst the positive qualifications of Kant's and (Neo-)Confucian philosophy, Palmquist's evaluation hits a crucial mark:

Western philosophy tends to define personhood in more abstract terms of the rights accorded to any human being simply by virtue of being human. Kant actually talks a great deal about both duties and rights; but he clearly gives priority in his System to duty. He put himself in the minority among Western philosophers by arguing not only that rights are an epiphenomenon of duty, rather than vice versa, but also that “practical reason” has priority over “theoretical reason”. Both of these tendencies appeal to Chinese philosophers, because, quite simply, they are inherently “Chinese” tendencies. Comparisons of Confucian ethics and Kantian ethics have, consequently, served as the springboard for much cross-cultural dialogue, especially from the Chinese side.<sup>4</sup>

---

1 See the theme edition on Kant and Chinese philosophy in the *Journal of Chinese Philosophy* 33.1 (2006), especially the following: Martin Müller, “Aspects of the Chinese Reception of Kant” (141–57); Gregory M. Reihman, “Categorically Denied: Kant's Criticism of Chinese Philosophy” (51–65); Wing-Cheuk Chan, “Mou Zongsan's Transformation of Kant's Philosophy” (125–39).

2 Stephen Palmquist, “How ‘Chinese’ was Kant?”, *The Philosopher* LXXXIV.1 (1996), 3–9; quoted from the abstract.

3 Cf. Julia Ching, “Chinese Ethics and Kant”, *Philosophy East and West* 28 (1978), 161–72; see also the authors in notes 1 and 2.

4 Palmquist, “How ‘Chinese’ was Kant?”

These and other comparative tendencies between Confucian and Kantian ethics have inspired an intense dialogue between Chinese and Western philosophers. However, discrepancies and dissimilarities between Confucian and Kantian way of philosophizing have been stressed as well. Various authors have noted that Confucian philosophy lacks epistemology and metaphysics as well as formal and fundamental concepts of moral agency, aspects that are of prime importance to Kant's philosophy.<sup>5</sup>

One of the most provocative asymmetries has been initiated by Kant himself, stating in his physical geography lecture notes: "Confucius teaches in his writings nothing outside a moral doctrine designed for princes ... and offers examples of former Chinese princes ... But a concept of virtue and morality never entered the heads of the Chinese."<sup>6</sup> Thus, Kant vehemently contrasts Confucius' way of thinking with his own philosophy. Kant's harsh way of opposing Chinese philosophy has evoked various authors to put forward religious, metaphysical, epistemological, and ethical interpretations. Some authors note the difference in fundamental religious concepts of God, Heaven, and future life. In this context quotations are drawn from Kant's *Physical Geography* to stress Kant's view that the Chinese did not believe in God or, if so, they make not much work out of it, that they have ill-defined notions of God, and that the Chinese "view the future life with the utmost indifference".<sup>7</sup> Kant's view is opposite to particularly European philosophers such as Wolf, Leibniz, Bayle, and Voltaire, who admired the sophisticated moral theory of Chinese philosophy.<sup>8</sup> Other authors have noted major differences in concepts of knowledge, experience, and science.

To address all these various interpretations is outside the context of this essay. My major aim is to argue that the above-quoted passage from Kant's physical geography lecture notes can, or perhaps, should be read from the perspective of Kant as a political philosopher, following the

---

5 Reihman, "Categorically Denied".

6 Quoted from Reihman, 58. This is Reihman's own translation from Kant's German text presented by Helmut von Glasenapp (ed.), *Kant und die Religionen des Ostens, Beihefte zum Jahrbuch der Albertus-Universität* (Kitzinger-Mainz: Holzner-Verlag, 1954), 103.

7 Reihman, 59.

8 Julia Ching and Willard G. Oxtoby, *Moral Enlightenment—Leibniz and Wolff on China* (Nettetal: Steyler Verlag, 1992).



approach of Reiss.<sup>9</sup> In this essay I will read Kant as a pragmatist, political philosopher, addressing the topic of personhood, morality, and power in politics and the ruling class. Understanding that Kant was arguing against a Machiavellian kind of politics and political philosophy may shed light on what he was opposing against certain strands in Chinese philosophy.

I will structure my argument as follows. First, the concerned quotation will be discussed in the context of Kant's philosophy. Second, the quotation will be placed in the setting of German philosophy, particularly Leibniz and Wolff. Third, the reception of Confucian philosophy by Wolff will be presented. Then Kant's basic critique on Chinese philosophy will be presented in terms of his view on moral and political agency. Finally, Kant's political-philosophical view on the role of rulers and their executives in political and societal life will be described.

## 2. Kant's Quotation in the Context of His Philosophy

The quotation comes from a manuscript containing Kant's notes for the physical geography lectures.<sup>10</sup> These lectures followed the prime interest in European thought regarding culture, religion, politics, and daily life in the Far East (e.g., India, Tibet, China, and Japan). Kant started lecturing on physical geography, or as Glasenapp claims how Kant phrased it, on this "useful and pleasant science", in the summer of 1756 and ended in the summer semester of 1796, leaving behind "lecture books" of 40 years of scholarly teaching.<sup>11</sup> He also prepared many lecture notes that made further observations, corrections, and additional re-

---

9 H. S. Reiss, *Kant's Political Writings*, tr. H. B. Nisbet, second edition (Cambridge: Cambridge University Press, 1991; here the 15th printing [2004]) of this 1991'edition has been used.

10 See Reihman; I also consulted the German version in volume 9 of the *Akademie Ausgabe* of Kant's *Werke*.

11 In German: "nützliche und angenehme Wissenschaft" (von Glasenapp, 3). Parts of these lecture books (in German: "Heften") have been published (as *Diktat-text*). The lecture notes or "Kollegnagschriften" (von Glasenapp, xi) have been collected by von Glasenapp, who selected passages for his publication in 1954. For the history of the publication(s) of Kant's lectures and lecture notes, see von Glasenapp. Kant was in his time considered as an excellent teacher; e.g., Herder's description of Kant's skills and talent as a teacher; students, but also colleague professors and even Prussian soldiers came to attend his lectures (Reiss, 193).

marks, as well as philosophical appraisals and evaluations. The concerned quotation is from these lecture notes.

It is interesting to note that the physical geography lectures as published by the Prussian Academy of Sciences contain detailed descriptions of the habits, conventions, and practicalities of Oriental life, ranging from systems of irrigation, the structure of the Great Chinese wall, specifics of food and consumptions, the nature of clothing, making compliments, science and language, and religion.<sup>12</sup> In some cases Kant uses evaluative or normative terms in these descriptions, but the passages themselves are nonetheless quite neutral. Thus, Kant wrote: "In China everybody has the freedom to throw away children who become a burden, through hanging or drowning."<sup>13</sup> Similarly, he describes how Chinese businessmen play their art as deceitful:

They deceive artfully. They can sew together a torn up piece of silk so nicely that even the most attentive businessman does not notice, and patch broken porcelain with an in-laying of copper wire so well that no one notices the initial crack. He is not ashamed if he is confounded in a lie, but is ashamed only to the extent that, through a slip he allowed the lie to be discovered.<sup>14</sup>

In contrast to such negatively felt and expressed references, Kant also notes the sophisticated and elevated character of many habits and customs in China, such as being kind, making compliments, talking softly, and no cursing or shouting.

The explicit and fierce normative assessments made by Kant emanate from his lectures notes or other sources. In these assessments Kant is making deep philosophical points. Thus, the quotation referred to above is not the result of racial or cultural bias, but addresses a fundamental philosophical problem. Reihman, Ching, and other authors are right in posing that deep philosophical issues are at stake.

---

12 There has been much discussion on racial and philosophical bias against the Orient in Kant's philosophy (see Reihman), but this is not the point here of this essay.

13 Reihman, note 22. See also page 57, where a similar comment is placed in context: "In Africa, theft is allowed, in China parents are permitted to throw their children on the street, the Eskimos strangle them, and in Brazil they are buried alive."

14 Reihman, 63 (note 18); Glasenapp, 83.

### 3. Kant and Philosophy on China in the 18th Century

Kant was never in China. In fact, he was never very far outside Königsberg. Therefore, he must have relied all his life on sources: philosophical books and manuscripts, traveler's books, and literature. Ching and Oxtoby state:

As a matter of fact, Kant did give lectures on Chinese moral philosophy and religion, all within the framework of his physical geography courses. The missionaries' translations of the classics must have been available to him. So also must have been the largely appreciative writings on China by his philosophical predecessors Leibniz and Wolff. Kant's lectures, however, reveal a close dependence on certain later publications, including travelers' reports. He was a man of his time—and his time had turned away from the earlier, heated interest in China and things Chinese. He dismissed as inspired propaganda the favorable reports of the Jesuit missionaries, the sources Leibniz and Wolff had used with so much enthusiasm.<sup>15</sup>

Thus, we have to position Kant in this Prussian (and wider, European) tradition of interest for China, that he, amongst others, reversed from a positive appraisal by philosophers such as Leibniz and Wolff to a negative judgment that was carried on in German and Prussian philosophy thereafter, such as in Hegel's philosophy.<sup>16</sup>

The crucial point here is not the discontinuity in appraisal of Chinese philosophy (i. e., the turn from positive thinking by Wolff and Leibniz to negative thinking by Kant and Hegel), but the continuity in terms of politics and political philosophy. The Jesuits, Leibniz, Wolff, but also Kant and Hegel discussed Confucianism as a political and political-philosophical issue.<sup>17</sup> Although the reception of Chinese philosophy in Prussia and Europe changed, from positive to negative, politics and political philosophy remained important issues, also in Kant's and Hegel's philosophy.

---

15 Ching and Oxtoby, 222.

16 See Ching and Oxtoby, particularly their Epilogue, "A reversal of opinion: Kant and Hegel on China", 221–9.

17 Ching and Oxtoby, 44.

#### 4. Wolff and the Reception of Confucian Philosophy

In the eighteenth century no school on politics and political philosophy was dominant in Germany and Prussia.<sup>18</sup> There was the school of Natural Law—Althusius, Grotius, and Pufendorf. Frederick the Great was influential, Leibniz as well, as were the cameralists. However, Wolff can be characterized as the leading philosopher of the “Aufklärung”, of the Enlightenment movement.<sup>19</sup> To understand better the particularities of Kant's response to Chinese philosophy it is worthwhile to elaborate on Wolff's views on China and Chinese philosophy. Wolff wrote extensive treatises on Chinese philosophy and Confucius, particularly in a political context.<sup>20</sup> He saw China as a role model, a country and a system of thinking, devoid of religion, that nevertheless had succeeded in perfecting spirituality, humanity, duty, and moral and political behavior.

In a nutshell Wolff's view on Chinese philosophy can be sketched as follows. In his text “Discourse on the Practical Philosophy of the Chinese”, Wolff starts by celebrating the fame of Chinese government in the past: “Since the earliest times, the wisdom of the Chinese has always been celebrated. No little praise has been lavished upon it, and upon the excellent constitution of Chinese government”.<sup>21</sup> Thus, Confucius was not the founder of Chinese philosophy, but the kings of ancient times:

Indeed, long before Confucius, the Chinese state distinguished itself through its excellent laws. Its princes gave their subjects a rule of the highest perfection, by their words as well as by their examples .... In fact, princes and subjects competed with each other in the practice of virtue. The ancient Chinese rulers and kings were also philosophers. Need we be surprised, consequently, if China experienced the truth of Plato's words that the state is happy where philosophers rule or where kings are philosophers?<sup>22</sup>

The first brilliant king was referred to as “Fo hi [Fu-hsi]”, who conceived the whole empire as one family, claiming that the same relationship between father and son should prevail between princes and sub-

---

18 Reiss, 9–10.

19 Reiss, 10.

20 See the lectures “Discourse on the Practical Philosophy of the Chinese” (text 1721 and notes 1726), in Ching and Oxtoby, 149–86 (plus the 1726 Preface by Wolff, 145–8) and “On the Philosopher King and the Ruling Philosopher”, in Ching and Oxtoby, 187–218.

21 Wolff, “Discourse”, 149.

22 Wolff, “Discourse”, 150.

jects.<sup>23</sup> Thus, a theory of social and political perfection, fidelity and virtue, and benevolent autocracy was put in place, later recovered by Confucius as the restorer of a “Fallen China and a state in collapse”.<sup>24</sup> Wolff saw much similarity between Confucian and his own philosophy, paying attention to the intellectual and moral qualities of rulers and their assistants:

Should it be asked what qualifications are requisite in a ruler to make a people happy, the answer is easy, as I have already demonstrated in my Politics: namely, a person capable of understanding the things requisite to render a people happy, as well as possessing a firm resolution for executing with fidelity what he knows to be conducive to so noble end.<sup>25</sup>

We will see below that Kant fundamentally disagreed with Wolff’s, and consequently also Confucius’,<sup>26</sup> view that the king was the first servant of the state, that the state should be run in the light of benevolent autocracy, and that politics is a virtuous exercise in statecraft only.

## 5. Kant’s Criticism on Wolff and Confucius

Kant’s criticism on Wolff’s and Confucius’ philosophy can be easily interpreted in terms of his practical *Critique*, particularly his views on moral and political agency. I will here follow the accounts of Allen W. Wood and Henry Allison, giving these accounts a political-philosophical twist in the next section.

In the *Groundwork* Kant contrasts an ethics of autonomy, whereby the will (*Wille*, or practical reason itself) is the basis of its own law, from the ethics of heteronomy, whereby something independent of the will, such as happiness, is the basis of moral law (4:440–1). Any “external” motive or command, be it goodness, benevolence, happiness,

---

23 Wolff, “Discourse”, 151, note 7.

24 Wolff, “Discourse”, 151–3.

25 See his “On the Philosopher King and the Ruling philosopher”i; the quotation comes from Ching and Oxtoby, 189.

26 It is not the claim of this essay that Confucian philosophy, let alone Chinese philosophy with its various schools and strands of thinking, are correctly represented here. However, Confucius saw the function of his ethics to be the cultivation of each individual person’s moral character, the wise governance of the country, and the regulation of the world. See Adrian Hsai, “Richard Wilhelm’s Reception of Confucianism in Comparison with James Legge’s and Max Weber’s”, *Journal of Ecumenical Studies* (2003), 40.

fortune, God, cannot ground autonomy as the basis of human dignity. Following Allison's incorporation thesis, two things are implied.<sup>27</sup> First, inclinations, desires, or other motivational forces can only constitute a reason for acting if the individual has incorporated these into his or her maxims, this being an act of spontaneity on the part of the agent. Second, these incentives or desires can provide maxims for personal conduct, when they are "taken as"—that is, "as mine" qua rational agent—only if they are to provide (grounded) motives or reasons for acting. Kant regards "the capacity to act on the basis of imperatives in general (not merely the categorical imperative) as the defining characteristic of free agency."<sup>28</sup> In this spirit Wood argues that commands of external origin—whether they come from God, kings, or any external source—cannot "provide a satisfactory account of moral obligation. Kant distinguishes the legislator of a law, the one who issues a command and may attach positive or negative sanctions to it, from the law's author, the one whose will imposes the obligation to obey it".<sup>29</sup>

In these terms, Kant has no objection to regarding the will of the rulers or any other social authority, be they ancestors, parents, families, communities, or spiritual or religious authority, as the *legislator* of the moral law. The gist of Kant's view is that only the rational will of the person obligated can be its *author*. Since Confucius is claimed by Wolff to have restored Chinese philosophy and government by addressing the precepts of the famous kings and rulers in ancient China, and since Kant is precisely opposing this stance in his quotation, my view is that we have to take a more radical path in this analysis, as attempted in the next section.<sup>30</sup>

---

27 H. E. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), 35–41; esp. 40.

28 Allison, 36.

29 Allen W. Wood, *Kant's Ethical Thought* (Cambridge: Cambridge University Press, 1999), 160–1. Wood discusses this in the context of theological morality, also addressing theories of perfection, such as to be found not only in Wolff and Stoic philosophy, but also in Confucius as expounded by Wolff and other authors in the eighteenth century who were well known to Kant.

30 The very idea to grasp back from the ancient kings for "perfection" is opposite to Kant's view, as are ideas on convention, prudence, craft, or virtue *simpliciter*. See also Kant's phrases in *Perpetual Peace* (Reiss, 231–2), where he says China has so much restriction on human freedom because it has no "external" enemy!

## 6. Kant as a Political Philosopher

The theme in Kant's political philosophy was to provide "a philosophical vindication of representative constitutional government, a vindication which would guarantee respect for the political rights of individuals."<sup>31</sup> Indeed, this is an important strand of thinking present in Kant's work: autonomy and freedom in social and political life. For this reason Reiss rightly calls him the "champion of liberalism".<sup>32</sup> However, there is a second strand of thinking in Kant's work: his Critical stance on the role of rulers and their executors as guardians of societal life. This strand of thinking has been largely neglected in the literature.<sup>33</sup> My claim will be that by filling up this lacuna we will also better understand Kant's harsh judgment on Confucius and Chinese philosophy.

Let us repeat the quotation by Kant in his physical geography lectures: "Confucius teaches in his writings nothing outside a moral doctrine designed for princes ... and offers examples of former Chinese princes ... But a concept of virtue and morality never entered the heads of the Chinese." Authors have focused on the last sentence: the claimed lack of a concept of virtue and morality, but not on the premise, that "Confucius teaches in his writings nothing outside a moral doctrine designed for princes ...". Now, the question is not whether Confucius has been correctly interpreted by Kant; this may or may not be the case. The issue of this essay is: what could Kant possibly have meant by stating this? What is wrong with designing a moral doctrine for princes? Particularly so, because it is said that Confucius was grappling back to the perfection of institutionalized government and moral behavior of princes and kings who virtuously and morally ruled Ancient China.

In his 1784 political essay "What is Enlightenment" Kant extensively argues the role of the rulers and the executors in society. There he starts (8:(54)):

*Enlightenment is man's emergence from his self-incurred immaturity. Immaturity is the inability to use one's own understanding without the guidance of another. This immaturity is self-incurred if its cause is not lack of understand-*

---

31 Reiss, 4.

32 Reiss, Postscript, 257.

33 See Rein Vos, "Public Use of Reason in Kant's Philosophy: Deliberative or Reflective?", in Valerie Rohden, Ricardo R. Terra, Guido A. de Almeida and Margit Ruffing (eds.), *Recht und Frieden in der Philosophie Kants—Akten des X. Internationalen Kant-Kongresses*, Band 4 (Berlin-New York: Walter de Gruyter, 2008), 753–63.

ing, but lack of resolution and courage to use it without the guidance of another. The motto of enlightenment is therefore: *Sapere aude!* Have courage to use your *own* understanding.

The only escape from immaturity is the freedom to use of one's own understanding. But then Kant states (8:(55)): "But I hear from all sides the cry: *don't argue!* The officer says: 'Don't argue, but rather march!' The tax official: 'Don't argue, pay!' The clergyman: 'Don't argue, believe! ... All this means restrictions on freedom everywhere.'"

Kant is arguing that, although it is right to restrict freedom of officers of the state and government to a certain extent—that is, these officers have to play their role—nevertheless, their freedom to use their own understanding should not be disregarded. Thus, Kant's cry out for *Sapere Aude (Dare to Think)*—that is *Have the Courage to Use Your Own Understanding*—holds out for public and governmental officers as well.<sup>34</sup> The problem Kant addresses here is how public officials, governed by authority and obedience, whence performing heteronomous actions, do not need to end up in quasi-mechanistic, passive obedience, but can remain responsible, self-determined professionals serving public ends.<sup>35</sup>

Kant opposes here the idea that the official duty of state officers is to surrender their judgment and responsibility to their superiors and their rulers, their kings or emperors. On the contrary, he exposes the idea that doctrines, laws and policies might be vague, contradictory, indeterminate and that these require interpretation and explication. An active servant of the government must exercise discretion, as any official should do in the spirit of his or her public office (serving public ends). The servant is not merely performing a technical task, a craft or prudence, but has to take up responsibility, judgment and morality.

Kant extends his portrayal of the servant to the "rulers" themselves. The head of the state, the ruler, "realises that there is no danger even to his *legislation* if he allows his subjects to make *public* use of their own reason and to put before the public their thoughts on better ways of drawing up laws, even if this entails forthright criticism of the current legis-

---

34 Kant's quotation has to be understood in terms of his opposite terms, "public use of reason" and "private use of private reason"; the latter is quite specific. What he has in mind is not that private reasoning is merely personal or individual, but that it is restricted to the "civil" post or office. See for further elaboration my paper, "Public Use of Reason".

35 Cf. Vos, "Public Use of Reason".



lation” (8:(59)). In this respect Kant rejects the philosophical view of Machiavelli—politics *is* power—‘arguing that politics and morality can be and should be united. In Appendix I of *Perpetual Peace* Kant states there are two kinds of politicians: the moral politician, “i. e., someone who conceives of the principles of political expediency in such a way that they can co-exist with morality”, and the political moralist, “i. e., one who fashions his morality to suit his own advantage as a statesman” (8:(118)).<sup>36</sup> The political moralist considers politics as a matter of prudence and expediency only, thus treats “the problems of political, international and cosmopolitan right as mere *technical* tasks” (8:(122)). The maxims of these politicians are: “Act first and justify your actions later”; “If you are the perpetrator, deny it”; and “Divide and rule.”<sup>37</sup> Kant, however, forcefully argues against the political moralist in favor of the moral politician. Thus, the problem of politics and morality directly relates to Kant’s problem of responsible, self-determined public officials in “What is Enlightenment”. Public officials are not merely part of a machine, not merely performing technical tasks, not merely exercising prudential skills; they have to become responsible, discerning, and moral professionals.

In this way we can better understand Kant’s quotation. Let us first consider the premise that “Confucius teaches in his writings nothing outside a moral doctrine designed for princes (and offering examples of former Chinese princes)”. On Kant’s view it is impossible to prescribe a moral doctrine. Any moral system has to be taken up as “an ought” by the agent, interpreting that system and critically reflecting on its vagueness, inconsistencies and faults. Only then can one start to possess a “concept of virtue and morality”. Since Confucius, according to Kant, teaches nothing outside a doctrine, he has no concept of virtue and morality, the conclusion that follows in Kant’s quotation.

## 7. Discussion and Conclusion

Against the background of the eighteenth century revolutions in France and the United States and the Enlightenment movement, Kant developed a Critical political philosophy. Kant’s aim was to arrive at philosophical principles that could be the basis for a just and lasting internal

---

<sup>36</sup> *Perpetual Peace*, Appendix I, 118.

<sup>37</sup> 8:(120). In Latin: “Fac et excusa”, “Si fecisti, nega”, and “Divide et impera”.

order and world peace. Kant defended the view that politics could be subjected to rational scrutiny and that accordingly political arrangements and institutions could be organized based on autonomy and freedom. Kant was an anti-authoritarian and a proponent of the "Rechtsstaat", the state governed according to the rule of law; but he defended this position in terms of human freedom in an age when "benevolent dynastic despotism was the prevailing mode of government."<sup>38</sup>

This critical stance toward rulers and their officers led Kant to his strong criticism of Confucius and Chinese philosophy. Kant rejected the political metaphor of the state as a family and of the relationship of the government and its citizens as the relationship of father and son. As Kant rejected patriarchic ways of thinking, he also rejected external sources of commands, such as benevolence, fidelity, perfection, and prudence. This Kantian criticism on Confucian philosophy was later taken up by Hegel:

The Chinese look on their moral rules as if they were laws of nature, positive external commandments, coercive rights and duties, or rules of courtesy. Freedom, through which the substantial determinations of reason alone can be translated into ethical attitudes, is absent; morality is a political matter which is administered by government officials and courts of law.<sup>39</sup>

Kant's criticism on Confucius and Chinese philosophy may seem (too) harsh. However, we should not forget that Kant's quotation came from his lecture notes, where he might have struck an informal tone. Moreover, and more importantly, Kant responded in a similar, condemning way to other philosophers, from the Stoics and Plato to Leibniz and Wolff, and even Herder, his own student. Thus, Confucian and Chinese philosophy are no separate chapter of Kant's criticism; they are in good company, so to speak.

Finally, we may ask whether Kant's criticism of Confucius and more broadly, of Chinese philosophy is right. It falls outside the context of this essay to establish whether Kant's reception and interpretation of Confucius and Chinese philosophy are correct. Obviously more in-depth study of Kant's works and the political and philosophical context in Germany, Prussia, and Europe in his time is necessary to establish the conclusions of this essay. In this respect, my analysis has a tentative char-

---

38 Reiss, 11.

39 Quoted from Ching and Oxtoby, 224–5. The authors' quotation comes from G. W. F. Hegel, *Lectures on the Philosophy of World History*, tr. H. B. Nisbet (Cambridge: Cambridge University Press, 1975), 145.

acter. But certainly Kant's Critical view on the freedom of rulers and public officials was adequate for his time and remains actual—the more so, one could say, in present times of a rapidly growing global world order. Kant's *Perpetual Peace* forecasted this global world where all countries would become connected and interdependent. However, in such a global world freedom, agency, and morality are even more necessary to improve government and policy and to find a moral basis for power. Kant's critical remarks on Confucius and Chinese philosophy were not uttered on the basis of cultural or racial bias. His criticism applied to politics in Europe and the United States and other countries as well, but particularly in his own Prussia. Kant had an open mind for the social, economical, cultural, and political conditions of each country and each region—as he acknowledged the political state of affairs of Germany and Prussia quite distinct from that in, for example, Britain and France. On Kant's view politics is a process whereby every country and region has to find its own path, but should not get off the track of integrating morality and politics in a genuine way. The basis of Kantian, moral politics is not doing good but doing right.

## 61. The Exclusion of Asia and Africa from the History of Philosophy: Is Kant Responsible?

Peter K. J. Park

As Kantianism was gaining ground in Germany in the last decade of the eighteenth century, Kantian philosophers began to rewrite the history of philosophy in a coordinated campaign to remake philosophy as discipline and knowledge. The Kantian Wilhelm Tennemann was among the first to exclude Asia from the history of philosophy and begin his account with the Greeks. To what extent was this exclusion the consequence of Kant's philosophy? To what extent was it the result of impersonal forces, specifically the *Verwissenschaftlichung* of philosophy and other university disciplines?<sup>1</sup> Could the exclusion of Asia, together with Africa, and modern philosophy's nascent Eurocentrism have been the result of racism?

Kant never lectured on the history of philosophy. He never produced a formal work of history of philosophy. He never published theoretical views on the history of philosophy, but the converts to his philosophy published in this area in his stead and sometimes with his approval. In their submissions to the Berlin Academy's prize-essay competition on the progress of metaphysics, Karl Leonhard Reinhold, Johann Heinrich Abicht, and Christian Friedrich Jensch argued that Kant's philosophy was a decisive step forward from Leibniz's and Wolff's. Kant's own response to the Academy's question survived into the nineteenth century.<sup>2</sup> "Lose Blätter zu den Fortschritten der Metaphysik", edited for *Kants gesammelte Schriften*, includes this fragment "on a philosophical [*philosophirende*] history of philosophy":

- 
- 1 There is no exact English equivalent to the German word *Verwissenschaftlichung*. Its approximate meaning is: transformation into a science or discipline.
  - 2 These manuscripts and related fragments were edited by Friedrich Theodor Rink as *Immanuel Kant über die von der Königl. Akademie der Wissenschaften zu Berlin für das Jahr 1791 ausgesetzte Preisfrage: Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnitzens und Wolf's Zeiten in Deutschland gemacht hat?* (Königsberg, 1804) in 20:255–351. The manuscripts have since been lost.

All historical knowledge is empirical and thus knowledge of things as they are, not as they must necessarily be ... An historical account of philosophy relates, how and in what order one has philosophized until now. However, to philosophize is a gradual development of human reason, and this could not have gone on or have even begun empirically, but, indeed, by concepts only. What reason compelled through its verdicts on things ... must have been a (theoretical or practical) need of reason to climb down to the ground [of things] and further to the first ground; from the very beginning through common reason ...<sup>3</sup>

Unlike ordinary history, the history of philosophy is not empirical. It is not characterized by chance and accident. The history of philosophy, as “a gradual development of human reason”, follows a logical necessity. Kant continues:

A philosophical history of philosophy is itself not historically or empirically possible, but rationally, that is *a priori* possible. For when it selects the *facta* of reason, it does not borrow them from historical narrative, but draws them from the nature of human reason; as philosophical archaeology.<sup>4</sup>

For Kant, the terms “historical” and “empirical” do not describe the work of the historian of philosophy.<sup>5</sup> So different is the history of philosophy from history that Kant suggested to rename it “philosophical archaeology.”

“How is an *a priori* history possible?” Kant posed the question in 1794 and alluded to the traits of a prophet (7:79–80). In a letter to Carl Morgenstern of August 14, 1795 (12:36), Kant flatters his friend, writing that he is capable “of composing a history of philosophy that does not follow the chronological order of books relating to it, but the natural order of the ideas which must successively develop themselves according to human reason.” In “Lose Blätter” (20:343) Kant describes the history of philosophy as “so special a kind that nothing of what is recounted therein could happen without knowing beforehand what should have happened and therefore also what can happen.” Thus, Kant himself seems to have prescribed the *a priori* construction of the history of philosophy.<sup>6</sup>

3 “Lose Blätter zu den Fortschritten der Metaphysik”, in 20:333–51. The passage appears on pages 340–41.

4 20:341. Kant elsewhere demotes “empirical history” as not scientific.

5 Cf. Giovanni Santinello (ed.), *Storia delle storie generali della filosofia*, vol.3: *Il secondo illuminismo e l'età Kantiana* (Padova: Editrice Antenore, 1988), Part II, 940–6.

6 Cf. Santinello, *Il secondo illuminismo*, Part II, 947–52.

While there were a half-dozen Kantians who contributed to the theory of history of philosophy, there were just two Kantians who actually dedicated labor to writing full-length histories of philosophy: Johann Gottlieb Gerhard Buhle (1763–1821) and Wilhelm Gottlieb Tennemann (1761–1819). The latter was a professor of philosophy at Jena before assuming Dieterich Tiedemann's chair of philosophy at Marburg in 1804. Tennemann produced the lengthiest history of philosophy written in a Kantian mode: the eleven-volume, unfinished *History of Philosophy* (1798–1819).<sup>7</sup> He also published a single-volume version of his history of philosophy, titled *Outline of the History of Philosophy* (1812, with four later editions), for students' use.<sup>8</sup>

In the introduction to the *History of Philosophy* Tennemann detailed the flaws of previous histories of philosophy: they were mainly collections of reports on the lives and opinions of philosophers. They made incomplete use of sources and used inappropriate sources. Previous histories of philosophy were poorly organized and lacked an overall plan.<sup>9</sup> They were simply copied out of earlier works “without critique, taste, discriminations” and “without philosophical spirit.”<sup>10</sup> They perpetuated “a mass of historical errors” and prejudices of the Church Fathers, who relied unfortunately on revelation and were biased in favor of the Jews. Subsequent historians of philosophy, the majority of them theologians, introduced the dubious notion of “antediluvian” philosophy and theological polemics into the history of philosophy. In sum, Tennemann regarded most previous histories of philosophy as unphilosophical compilations and chronicles.<sup>11</sup>

Tennemann contended that while the history of philosophy shared certain characteristics with other genres of history, it was a genre separate from the history of nations, the history of scholarship, and the history of other sciences. He even cautioned against confusing the history of the literature of philosophy with the history of philosophy proper.<sup>12</sup>

---

7 Wilhelm Gottlieb Tennemann, *Geschichte der Philosophie*, 11 vols. (Leipzig: Johann Ambrosius Barth, 1798–1819).

8 Wilhelm Gottlieb Tennemann, *Grundriß der Geschichte der Philosophie für den akademischen Unterricht* (Leipzig: Johann Ambrosius Barth, 1812; later editions of 1816, 1820, 1825, and 1829).

9 Tennemann, *Geschichte der Philosophie* 1:IV.

10 Tennemann, *Geschichte der Philosophie* 1:III and LXXIV.

11 Tennemann, *Geschichte der Philosophie* 1:LXXIV–LXXXV. Pierre Bayle is Tennemann's one exception.

12 Tennemann, *Geschichte der Philosophie* 1:XLVI.

In his *Outline of the History of Philosophy* Tennemann differentiated the history of philosophy from the history of mankind, intellectual history, history of sciences, biography, literary history, analysis of works, and compilations of opinions.<sup>13</sup> Tennemann cautioned however that the history of philosophy should not be so exclusively concerned with the exposition of philosophical systems that it omits the historical dimension.<sup>14</sup>

Tennemann prescribed the kinds of sources to be used for the history of philosophy. Philosophical writings, literary works, reports and investigations by observers, and other historical data could be admitted. However, the “philosophemes” (the real objects of the history of philosophy) should be taken from the philosopher’s own writings. Extra-philosophical writings were to be treated as supplementary sources.<sup>15</sup>

Since the information from a now delimited fund of sources could not all be incorporated into a history, it was important to decide what was to be included and what was not. Tennemann laid out some rules: whatever has “a relation to and influence on the formation of this science [philosophy]” may be included, but whatever “disrupts the coherency and overview of the history” should not be. Detailed biographies of philosophers, for example, should not be included as these would “injure the unity of the history” and inappropriately connect the actual object of inquiry, philosophy, to the personal histories of the philosophers. Details from the life of a philosopher may still be woven into the history of philosophy, but only if doing so enhances a coherent view of philosophy’s development.<sup>16</sup>

Tennemann only then laid out some definitions. “Science [*Wissenschaft*]” is a “system of knowledge.” “[R]eason is the only source of all science; for every science is an architectonically rendered structure for which reason draws up the idea and guides the completion.” The “idea of science” was a “necessary expression of reason”, subsisting through all the changes of the science’s history.<sup>17</sup> Tennemann presented a definition of the history of philosophy (in Kantian terms): “History of philosophy is exposition of the successive development of philosophy or

---

13 Tennemann, *Grundriß* (1816), 7.

14 Tennemann, *Geschichte der Philosophie* 1:XLVI.

15 Tennemann, *Geschichte der Philosophie* 1:LVII.

16 Tennemann, *Geschichte der Philosophie* 1:LXIV–LXV.

17 Tennemann, *Geschichte der Philosophie* 1:XIV–XVI.

exposition of the exertions of reason to realize the idea of the science from the final grounds and laws of nature and freedom.”<sup>18</sup>

Historiographers seem to be unanimous that the history of philosophy in Germany underwent a fundamental reform in the late eighteenth century. I have argued elsewhere that it was the Kantians, starting in the 1790s, who were the central agents of this reform.<sup>19</sup> Philological methods that characterize modern historical practice, such as source criticism and writing history from authenticated primary sources, were already innovated and in practice in German academies and universities and were not the substance of the reforms that took place at the end of the eighteenth century. What was new about the end-of-century reforms was that the empirical record of the history of philosophy was now subordinated to an *a priori* organization, whose principles were derived from Kantian philosophy. Despite the blatant sectarianism of such an approach, historiographers have tended to see this development as a decisive moment in philosophy’s and the history of philosophy’s *Verwissenschaftlichung*. The rest of my essay takes up the following questions: Does *Verwissenschaftlichung* adequately explain the exclusion of non-European thought from the history of philosophy? Does Enlightenment racism better explain this exclusion? Is Kant responsible for this exclusion and for modern philosophy’s Eurocentrism?

It is not well known that Tennemann was among the first modern European authors to exclude Asia and Africa from the history of philosophy and begin his account with the Greeks. Historians of philosophy before Tennemann began their accounts with one or another ancient people of the Old Testament. In numerous examples of history of philosophy from the sixteenth, seventeenth, and eighteenth centuries, Adam or Moses (or the Egyptians or the Jews) figure as the first philosophers. As late as 1788, the Leibnizio-Wolffian philosopher Johann August Eberhard published a *General History of Philosophy* that treated first the philosophical ideas of “non-Greek peoples”: the Hebrews, Chaldeans, Persians, Arabs, Egyptians, Indians, Chinese, Phoenicians, Scythians, Getes, and Celts.<sup>20</sup> The first volume of Tennemann’s *History of Phi-*

18 Tennemann, *Geschichte der Philosophie* 1:XXIX.

19 See my dissertation, *The Exclusion of Asia from the Formation of a Modern Canon of Philosophy: Debates in German Philosophy 1790–1830* (University of California, Los Angeles, 2005), 9–46.

20 Johann August Eberhard, *Allgemeine Geschichte der Philosophie zum Gebrauch akademischer Vorlesungen*<sup>2</sup> (Halle: Hemmerde und Schwetschke, 1791 [1788]).



*losophy* begins right away with Greek philosophers. The author does not explain why the Orientals are excluded, and he does not acknowledge in any way that he was breaking radically from historiographical tradition. The first line of the first chapter reads, “The Greek nation is singular in history ...” and is followed by some arguments for Greek originality. It continues, “The physical and political constitution of Greece, the spirit and character, the education and activity of the inhabitants [brought] together so many important advantages for the development and cultivation of the human mind, which one will not easily come across in other countries of the time.” Greek philosophy met the requirement that philosophy be independent of political interests.<sup>21</sup> While stating that he does not deny the philosophical spirit in other nations, Tennemann asserted that the Greeks developed it independently. Philosophy developed in Greece “without the admixture of foreign elements” and was what was transmitted to all later civilized peoples.<sup>22</sup> The argument here is that Greek philosophy was an autochthonous development, that it is furthermore the only philosophy that matters since it was passed down to the Romans and then to the Germanic nations. Greek philosophy, not Oriental philosophy, was the first ancestor of modern philosophy.

That philosophy has Greek origins was an opinion held by only three historians of philosophy in the eighteenth century—all active during the late eighteenth century. Besides Tennemann, Christoph Meiners (1747–1810) and Dieterich Tiedemann (1748–1803) excluded the Orient from their histories of philosophy. Tiedemann tersely stated his reasons for the exclusion; these can be summed up by the following: the so-called philosophical ideas of the Chaldeans, Persians, Indians, and Egyptians “contain mere poetry of times still half-brutish” or are based on revelation.<sup>23</sup> He urged that such ideas be barred from the history of philosophy. Meiners, a teacher and friend of Tiedemann’s, elaborated his reasons over copious pages in several of his books and articles. He is the key to explaining the exclusion of Asia and Africa from the history of philosophy.

---

21 Tennemann, *Geschichte der Philosophie* 1:3, 5, 11. He adds, “It is true that here one finds some examples of the persecution and restriction of the freedom of thought, but this is true in respect to individual philosophers, not in respect to philosophy itself.”

22 Tennemann, *Geschichte der Philosophie* 1:3, 5.

23 Dieterich Tiedemann, *Geist der spekulativen Philosophie*, 6 vols. (Marburg: Akademische Buchhandlung, 1791–1797), 1:xviii.

Meiners was a professor of philosophy at Göttingen from 1772 to 1810 and the author of at least forty-four monographs (including several multi-volume works) and one-hundred-eighty journal articles in psychology, aesthetics, the histories of science, philosophy, and universities, and the natural history (ethnography) of ancient and modern peoples. Britta Rupp-Eisenreich has recognized that he played a significant role in shaping early anthropology in Germany and France through his numerous publications.<sup>24</sup> In 1781–82, he published a two-volume history of ancient Greek and Roman science that by the nineteenth century reached a Europe-wide readership through a French translation prepared by J. C. Laveaux.<sup>25</sup> Meiners is included in Johann Gustav Droysen's account of the "Göttingen Historical School" that Droysen credited with developing *Weltgeschichte*, *Universalgeschichte*, and the natural history of mankind, in Germany. Meiners was one of the main contributors to the Enlightenment science of man, "the central science of the time", "the royal science of the second half of the century."<sup>26</sup> Significantly, Meiners was a professor at the university in Göttingen, one of the centers of the German Enlightenment.<sup>27</sup> According to John Zammi-

---

24 Britta Rupp-Eisenreich, "Des choses occultes en histoire des sciences humaines: Le destin de la 'science nouvelle' de Christoph Meiners", *L'ethnographie* 90–91 (1983), 131–83.

25 Christoph Meiners, *Geschichte des Ursprungs, Fortgangs und Verfalls der Wissenschaften in Griechenland und Rom*, 2 vols. (Lemgo: Meyer, 1781–1782); tr. J. C. Laveaux as *Histoire de l'origine, des progrès et de la décadence des sciences dans la Grèce*, 4 vols. (Paris, an VII [1799]); Santinello, *Il secondo Illuminismo*, Part II, 725.

26 Hans-Jürgen Schings, "Der philosophische Arzt", in *Melancholie und Aufklärung: Melancholiker und ihre Kritiker in Erfahrungsseelenkunde und Literatur des 18. Jahrhunderts* (Stuttgart: Metzler, 1977), 13; Helmut Pfotenbauer, *Literarische Anthropologie* (Stuttgart: Metzler, 1987), 1; John Zammito, *Kant, Herder, and the Birth of Anthropology* (Chicago: University of Chicago Press, 2002), 245, 345. The best recent study of Meiners' race science is a 46-page essay by Friedrich Lotter, "Christoph Meiners und die Lehre von der unterschiedlichen Wertigkeit der Menschenrassen," in Hartmut Boockmann and Hermann Wellenreuther (eds.), *Geschichtswissenschaft in Göttingen: Eine Vorlesungsreihe* (Göttingen: Vandenhoeck and Ruprecht, 1987), 30–75.

27 Wilhelm Dilthey called Göttingen "the center of historical studies in Germany" (*Studien zur Geschichte des deutschen Geistes*, in Dilthey's *Gesammelte Schriften*, ed. P. Ritter [Stuttgart: Teubner, 1959], 3:261).

to, Meiners laid out a theory of knowledge in *Revision der Philosophie* (1772), a book that served as a manifesto for the Göttingen School.<sup>28</sup>

Among Meiners' numerous works is a one-volume textbook of the history of philosophy, *Outline of the History of World Wisdom [or Philosophy]* (1786; second ed., 1789).<sup>29</sup> It is the only book by Meiners in the history of philosophy genre. What Meiners argues about Oriental philosophy in this book is consistent with his account of the rise of (scientific) civilization in *History of the Origin, Progress, and Decline of the Sciences in Greece and Rome* (1781) and his overall vision of human nature in *Outline of the History of Mankind* (1785; second ed., 1793).<sup>30</sup> Bringing these works together and reading them in conjunction with Tennemann's history of philosophy will yield the ultimate explanation as to why Tennemann excluded Asia and Africa from the history of philosophy.

With alarmist rhetoric opening his Preface, Meiners censured the "unhistorical enthusiasm" spreading to "certain secret schools among us" (in Germany?) of attributing "the most groundless ideas and systems" to "the raw or little-cultured peoples of the oldest antiquity." He characterizes this "unhistorical enthusiasm" as an "illness of the mind", "the effect of the demise of all genuine scholarship and critique" suffered by the French ("our neighbors beyond the Rhine") as much as by the ancient Greeks and Romans!<sup>31</sup> In *History of the Origin, Progress, and Decline of the Sciences in Greece and Rome*, published five years earlier, Meiners had already begun his campaign against the traditional opinion that settlers from Asia and Africa transmitted their sciences and arts to the uncivilized, aboriginal Greeks. In that work, Meiners asserted that this was improbable and even contrary to history. From his reading of classical sources he deduced that these foreigners were more like "refugees" than settlers, driven out of their home country by the fear of punishment for crimes they had committed or by powerful opponents, and that they did not have time enough to prepare themselves for the long years required to found new cities in Europe or to bring along every

28 Christoph Meiners, *Revision der Philosophie* (Göttingen und Gotha: Dieterich, 1772); Zammito, *Kant, Herder, and the Birth of Anthropology*, 248.

29 Christoph Meiners, *Grundriß der Geschichte der Weltweisheit*<sup>2</sup> (Lemgo: Meyer, 1789 [1786]).

30 Christoph Meiners, *Geschichte des Ursprungs, Fortgangs und Verfalls der Wissenschaften in Griechenland und Rom*, 2 vols. (Lemgo: Meyer, 1781–1782) and *Grundriß der Geschichte der Menschheit*<sup>2</sup> (Lemgo: Meyer, 1793 [1785]).

31 Meiners, *Weltweisheit*, Preface.

kind of knowledge and useful article necessary for a civilized existence. According to Meiners, these refugees encountered in Greece “men who had an invincible hatred of foreigners”, who frequently attacked and robbed them, and who could not be driven out of their lands or become accustomed to the lifestyle of the newcomers. Even after they were well established in Greece and developed advanced weapons, they remained surrounded by numerous undeterred tribes who fought against them. It did not help the settlers to have maintained no contact with their countries of origin, for they were then “plundered and carried off as much as others by their former countrymen.” Egyptian and Phoenician settlers, not being able to subdue the Greeks, “handed down little or nothing of cultivating knowledge and skills, except for their gods and rituals of worship, the beginnings of agriculture, and a completely useless script, and a certain number of words.”<sup>32</sup>

Meiners’ thesis is that, despite Egyptian and Phoenician settlements, the native inhabitants of Greece did not inherit a civilization from either nation. Following Diogenes Laertius, Meiners claimed that philosophy originated in the Greek cities on the western edge of Asia (i.e., Ionia). Conditions in this region promoted material prosperity, and this in turn promoted the first flowering of the arts and sciences on earth.<sup>33</sup> According to Martin Bernal, playing down the role of Phoenicians and Egyptians in the founding of Greek civilization was central to the strategy of scholars, starting in the late eighteenth century, who reimagined a Greek pedigree for European civilization in conformity with their assumptions about European racial superiority.<sup>34</sup> I will presently show how well the evidence bears out Bernal’s claim.

A historiographer of philosophy who reads Meiners’ *Outline of the History of World Wisdom*, but does not read *Outline of the History of Mankind* and other anthropological works by Meiners, may never realize that the author’s account of the historical origins of philosophy are completely consistent with his racist anthropological views. The *Outline of the History of Mankind*, introducing students to Meiners’s anthropology, begins by laying out in a synoptic fashion the conclusions of Meiners’s anthropological research. In the opening pages Meiners presents the

---

32 Meiners, *Geschichte des Ursprungs*, 3–4.

33 Meiners, *Weltweisheit*, 24.

34 Martin Bernal, *Black Athena: The Afroasiatic Roots of Classical Civilization*, vol.1: *The Fabrication of Ancient Greece 1785–1985* (New Brunswick, New Jersey: Rutgers University Press, 1987), 189–443.

human species as two large divisions: “the Tartar, or Caucasian” and “the Mongolian”. Meiners listed under the second division the Chinese, Tibetans, Kalmucks, Samoyeds, lower-caste Hindus, and the blacks of New Guinea, New Holland (Australia), and Africa. The peoples of the Mongolian division are categorically described as “weaker in body and mind” and “more depraved and vicious” than those of the Caucasian division. Under the Caucasian division, Meiners distinguished two races (*Racen*): Celtic and Slavic. Under the Celtic race, he listed Greeks, Germans, Italians, Gauls, Spaniards, Britons, Irishmen, and Scandinavians. He did not provide any specific examples for the Slavic race. In Meiners’ scheme, the Celtic race possesses superior intellectual and moral qualities to those of the Slavic race. (Apparently, Meiners’ anthropological studies did not turn him into a cultural or moral relativist.) In the second edition of *Outline of the History of Mankind*, published in 1793, the names Caucasian and Mongolian are replaced with “the white, or light-colored, and beautiful” and “the dark-colored and ugly.”<sup>35</sup> This simplification was probably an attempt to eliminate confusion and criticism arising from his initial naming of the two large divisions. Meiners now also introduced a third white race. The Egyptians, Jews, Arabs, Persians, and upper-caste Hindus were brought together under an “Oriental” (white) race. In 1793, the Celtic peoples were still given the distinction of being the white race with the greatest intellectual and moral qualities.<sup>36</sup>

After reading Meiners in some depth, it becomes apparent that innate differences between the races explained for him literally everything about the course of human affairs, beginning with the way human groups were dispersed over the earth and the ancestry and kinship between nations. Racial differences explained why the “great law-givers, sages, and heroes” were white and why Mongolian peoples never developed sciences. If some dark-and-ugly nations did exhibit some scientific activity, this could have only come about through their interaction with Whites. In any case, history showed Meiners that the arts and sciences tended to degenerate in the hands of dark-and-ugly nations. Finally, racial differences explained why Europeans have almost always dominated

---

35 Meiners, *Menschheit* (1793), 5.

36 Meiners, *Menschheit* (1793), 29–30. The author adds (75), “All these white peoples have several common characteristics [cit. omitted] yet the Slavic and the Oriental peoples are in agreement with each other more than with the Germanic or other Celtic nations [cit. omitted].”

all other peoples of the earth and why political rights have existed among whites, while “the most horrible despotism slams its unshakeable throne upon the majority of peoples of the earth.” For Meiners, the laws and political constitutions of European nations and the European Enlightenment itself were direct evidence of the superior intellectual and moral faculties innate to their race.<sup>37</sup>

To the historian’s great fortune, Meiners left behind works of both ethnography and history of philosophy (and sciences). He allows us to see how racism and Eurocentrism in philosophy go hand-in-hand. Tennemann, who did not write ethnography and never discussed in print the differences between human races, deployed in his history of philosophy a set of (latently-racist) anthropological arguments that Meiners innovated and presented to the reading public, years earlier. Tennemann agreed with Meiners that philosophy was Greek in origin: the Greeks invented and developed it while all other peoples did not, and Greek philosophy was what was passed down to the Romans and the modern peoples of Europe. The successive stages or periods of this passing-down constitute the history of philosophy.<sup>38</sup> It so happens that the philosophers in this line of succession are all Celtic whites. If an Oriental nation had or has a philosophy, it came in possession of it by appropriating the learning of the Greeks or another white nation.

Tennemann’s arguments for the Greek origin of philosophy were copied right out of Meiners’s publications, specifically *History of the Origin, Progress, and Decline of the Sciences in Greece and Rome* and *Outline of the History of World Wisdom*.<sup>39</sup> The component arguments regarding, for example, climate, material prosperity, and political culture that make up Tennemann’s claim of the Greek origin of (scientific) civilization were conceived earlier in the Enlightenment, but Meiners was the one who brought them together in the form one finds them in Tennemann’s *Outline of the History of Philosophy*.

Earlier, I showed that Tennemann’s approach to the history of philosophy conforms to Kantian principles that determined the organization and content of that history. But is the exclusion of non-European thought from the history of philosophy also in conformity with Kant’s thought? In recent years, several important essays have appeared on the

---

37 Meiners, *Menschheit* (1793), 30–1.

38 Tennemann, *Grundriss* (1820), 10–1.

39 There are more than a couple of direct citations to Meiners in Tennemann’s *Grundriß der Geschichte der Philosophie*.

topic of Kant's racism. Two scholars in particular, Robert Bernasconi and Mark Larrimore, have done much to raise awareness (anew) of Kant's thinking on race and stimulate debate on whether and how Kant the racist can be reconciled with Kant the cosmopolitan and moral universalist.<sup>40</sup> Bernasconi and Larrimore have revealed a major European philosopher who, over the whole length of his teaching career, regularly returned to the problem of human diversity and inequality and sought to make positive sense of human affairs with a theory and anthropology of race.

Kant was a powerful thinker of race, who was not incapable of adopting the racist ideas of others. I argue that a kind of racist feedback-loop existed between Kant and Meiners. Frank W. P. Dougherty has already noted that Meiners incorporated Kant's definition of race into the second edition of *Outline of the History of Mankind* and even included an explicit reference to Kant's 1785 essay.<sup>41</sup> Kant shared more than a few racial-anthropological descriptions and opinions with Meiners. Because the latter published overwhelmingly more in empirical anthropology than the former, it is more likely that the former got his racial descriptions from the latter. In the transcription of Kant's anthropology lectures, the name Meiners does not appear. This is not surprising, given that Meiners was a strident critic of Kant and a rival in anthropology.

Kant's description of the races, like Meiners's, entailed a hierarchy of relative worth. Like Meiners in his article "On the Population of America" (1788), Kant in his anthropology lectures described the native Americans as uneducable, emotionally and sexually unexcitable, barely

---

40 Robert Bernasconi, "Who Invented the Concept of Race? Kant's Role in the Enlightenment Construction of Race", in Robert Bernasconi (ed.), *Race* (Malden, Mass.: Blackwell, 2001), 11–36; "Kant as an Unfamiliar Source of Racism", in Julie K. Ward and Tommy L. Lott (eds.), *Philosophers on Race: Critical Essays* (Malden, Mass.: Blackwell, 2002), 145–66; and "Will the Real Kant Please Stand Up: The Challenge of Enlightenment Racism to the Study of the History of Philosophy", in *Radical Philosophy* 117 (2003), 13–22; Mark Larrimore, "Sublime Waste: Kant on the Destiny of the 'Races'", in *Canadian Journal of Philosophy* 25, supp., *Civilization and Oppression* (1999), 99–125.

41 Frank W. P. Dougherty, "Christoph Meiners und Johann Friedrich Blumenbach im Streit um den Begriff der Menschenrasse", in Gunter Mann and Franz Dumont (eds.), *Die Natur des Menschen: Probleme der Physischen Anthropologie und Rassenkunde (1750–1850)* (Stuttgart and New York: Gustav Fischer, 1990), 102n52.

fertile, weak, and lazy.<sup>42</sup> Like Meiners in “On the Nature of the African Negroes” (1790), Kant described Blacks as being trainable for servitude, but not capable of self-governance and moral independence.<sup>43</sup> Meiners’ claim of a Greek origin of (scientific) civilization and his characterization of Oriental thought are consonant with Kant’s statement that the Hindu (or Indian or yellow) race never achieved an abstract concept and that their moral precepts are thus not based on principles. Kant said in a lecture (25.2:665): “All Oriental peoples are not in the position to establish through concepts a single property of morality or law. Rather, all their morals are based on appearances.” To be completely clear, yellow peoples are “not in a position”—meaning they lack the capacity—to produce science (or philosophy). Only white peoples have the capacity for abstract concepts. Also, the whites are the ones who brought about all the revolutions in history. The Hindus, Americans, and Negroes have never been agents of history. The white race is the only one marked by historical progress.<sup>44</sup>

Again, Kant taught that the Hindu (or yellow) race did not develop philosophy because they did not have that capacity. In his anthropology lectures, Kant explicitly attributes this lack *not* to the form of government or customs of the Orientals, but to their descent (*Abstammung*) (15.2:880). Others had claimed that the form of government or customs of a people determined its character. Kant taught his students that it was the other way around. It was racial character that determined the form of government and customs (15.2:880). I have discovered that Meiners presented the full version of this argument in his article “On the Causes of Despotism” (1788).<sup>45</sup> Here, Meiners argues that “weaknesses, lack of feeling, and limitation of mind” of Mongolian peoples led to their complete subjugation by their lords.<sup>46</sup> “A similar weakness, lack of feeling, and idiocy”, along with “cowardice”, in negroes and Americans were

42 25.2:1187–8; Meiners, “Ueber die Bevölkerung von America”, in Christoph Meiners and Ludwig Thimoteus Spittler (eds.), *Göttingisches historisches Magazin* 3 (Hannover: Helwing, 1788): 193–218.

43 25:877–9 (R1520); Meiners, “Ueber die Natur der Afrikanischen Neger”, in *Göttingisches historisches Magazin* 6 (1790), 385–456.

44 25.2:1187–8. Kant also said (25:877–9 [R1520]): “Von der race der Weissen, die alle revolutionen in der Welt hervorgebracht hat .... Die drey übrige racen gar keine.”

45 Meiners, “Ueber die Ursachen des Despotismus”, in *Göttingisches historisches Magazin* 2 (1788): 193–229.

46 Meiners, “Despotismus”, 203



the causes of their easy submission to arbitrary domination.<sup>47</sup> “Therefore, the freedom or slavery of peoples has existed in all parts of the earth and in all the times like the inner worth and unworth of the same, and never was a nation oppressed by a despot without deserving this destiny and having forged its own chains.”<sup>48</sup> The claim that despotism prevented philosophy’s development in the Orient is not really a cultural analysis. Already in Tennemann, it is a racist trope.

Kant’s views on race can also be found in his notes titled *Reflexionen*.<sup>49</sup> One note in particular, *Reflexion 1520*, further confirms my thesis. There, Kant affirms the white ancestry of Europeans: “Our (ancient) human history goes back with reliability only to the race of the Whites. Egyptians. Persians. Thracians. Greeks. Celts. Scythians. (not Indians, Negroes.)”<sup>50</sup> Kant denied that the ancestry and civilization of the whites went back to the “Indians” or “Negroes” (the non-white races of the Old World). For Kant, human history concerns the peoples of “Celtic stock”.<sup>51</sup> In Larrimore’s analysis of *Reflexion 1520*, Kant excluded the non-white races from human history because they made no contribution to that history. They made no contribution because, as Larrimore puts it, “they did not have it in them to do so.”<sup>52</sup> I draw the easy corollary that non-white peoples are excluded from the history of philosophy because they made no contribution to it because they did not have it in them to do so.

I believe that the thesis of *Verwissenschaftlichung* does not explain the process that led to the exclusion of Asia and Africa from the history of philosophy. I believe that racism better explains the phenomenon. I have shown that Tennemann’s exclusion of Asia and Africa was firmly grounded in the racial anthropology that Kant and Meiners constructed during the last quarter of the eighteenth century. Kant and Meiners were a tag-team, working in tandem to produce a distinctly modern science of race. Kant is as responsible as Meiners for the exclusion of Asia and Africa from the history of philosophy and for the rise of Eurocentrism in the modern discipline of philosophy.

---

47 Meiners, “Despotismus”, 208.

48 Meiners, “Despotismus”, 200.

49 15.2 (“Reflexionen zur Anthropologie”).

50 25.2:877–79: “Unsere (alte) Geschichte der Menschen geht mit Zuverlässigkeit nur auf die Race der Weissen. Egypter. Perser. Thracier. Griechen. Celten. Scythen. (nicht Indier, Neger.)”

51 25.2:880. The term “Celtischer Stamm” appears in *Reflexion 1520*.

52 Larrimore, “Sublime Waste”, 115.

## 62. Menschliche Autonomie als Aufgabe – der Autonomiebegriff in der Geschichtsphilosophie Kants

*Soo Bae Kim*

### 1. Mensch als „animal rationabile“

Wie bekannt ist, versucht Kant die Realität des Moralischen zu begründen, indem er das Bewusstsein des unbedingt gebietenden Moralgesetzes als „Faktum der Vernunft“ bezeichnet. Unser Bewusstsein des Moralgesetzes, so meint Kant, sei kein Hirngespinnst, sondern in „aller Menschenvernunft gewesen und ihrem Wesen einverleibt“.<sup>1</sup> Nach ihm ist der Mensch „als zur intelligibelen [Welt] gehörig nicht bloß ... *gedacht* ..., sondern sogar *in Ansehung des Gesetzes* ihrer [d.h. der Freiheit] Kausalität bestimmt und assertorisch *erkannt*“ (*KpV*, A188). „Ein jedes Wesen, das nicht anders als *unter der Idee der Freiheit* handeln kann, ist eben darum, in praktischer Rücksicht, wirklich frei, d. h. es gelten für *dasselbe* alle Gesetze, die mit der Freiheit unzertrennlich verbunden sind“.<sup>2</sup> „Er [d.h. der Mensch] ist nämlich das Subjekt des moralischen Gesetzes, welches heilig ist, vermöge der Autonomie seiner Freiheit“ (*KpV*, A156). Als moralisches Wesen betrachtet, ist der Mensch von den sinnlichen Bedingungen unabhängig und völlig autonom. Er ist sozusagen ein Mitglied der intelligiblen Welt und selber Ding an sich. Von der zeitlichen Beschränkung zu sprechen, egal ob es sich dabei um einen Fortschritt oder einen Rückschritt handelt, scheint daher für ihn sinnlos zu sein. Als das durch die praktische Vernunft begabte Wesen kann und soll er sowohl in der Vergangenheit als auch in der Gegenwart oder auch in der Zukunft dem von ihm selbst auferlegten Gesetz folgen.

---

1 *Kritik der praktischen Vernunft* (= *KpV*), A188. Kants Schriften werden nach der Ausgabe von Wilhelm Weischedel: *Immanuel Kant, Werke in sechs Bänden*,<sup>5</sup> Bd.4 (Darmstadt: Wissenschaftliche Buchgesellschaft, 1983 [1956–1964]) zitiert, und zwar nach der dort vermerkten Paginierung der Originalausgaben; A bezeichnet die erste, B die zweite Auflage.

2 *Grundlegung zur Metaphysik der Sitten* (= *GMS*), Bd.4, B100.

Im Unterschied zu der Ansicht in der *GMS* oder in der *KpV*, wo die Freiheit im praktischen Sinne und auch die praktische Vernunft von einem zeitlosen Standpunkt her verstanden wird, wird der menschliche Wille [oder die Willkür] in der *Kritik der reinen Vernunft* als „arbitrium sensitivum ... liberum“ bezeichnet.<sup>3</sup> Dieser Wille ist an sich noch nicht vollkommen freier Wille. Er ist weder der Wille desjenigen Menschen, dessen Handlungsursachen der Gegenstand der Erkenntnis werden können, noch der des Menschen als Ding an sich. Nicht weit entfernt von dieser Auffassung scheint es zu sein, wenn Kant in seiner Anthropologievorlesung den Menschen als „animal rationabile“ und nicht als „animal rationale“ charakterisiert. Nach dieser Charakterisierung, die eine phänomenologische Bestimmung des Menschen heißen könnte, ist die menschliche Autonomie „nicht eine Eigenschaft jedes beliebigen Individuums“, sondern eine „Aufgabe“.<sup>4</sup> Wenn es sich hier um keine fehlgehende Interpretation, sondern um bloße, getreue Rekonstruktion Kantischer Texte handelt, eröffnet es m. E. neue geschichtsphilosophische Horizonte, die in diesem Begriff enthalten sind.

## 2. „Zeitlosigkeit“ vs. „Geschichtlichkeit“

Kants geschichtsphilosophische Gedanken spiegeln das oben genannte phänomenologische Menschenbild wider. Geschichtliche Fakten zeigen nach Kant keine Regularität, geschweige denn irgendwelchen Sinn. Sie sind an sich nichts weiter als ein Durcheinander von „Tun und Lassen“, das „aus Torheit, kindischer Eitelkeit, oft auch kindischer Bosheit und Zerstörungssucht zusammengewebt“ ist.<sup>5</sup> Den Grund, warum es so sein muss, findet Kant in der Tatsache, dass sich die Menschen „nicht bloß instinktmäßig, wie Tiere, und doch auch nicht, wie vernünftige Weltbürger“ verhalten (*IaG*, A387). Anders gesagt ist der Mensch zwar mit dem Vermögen begabt, eine von seinen Handlungsoptionen frei zu wählen, aber neigt dennoch dazu, den sinnlichen Begierden nachzugeben. Der Prozess der menschlichen Geschichte kann also weder dem mechanistisch festgesetzten Gesetz gemäß verlaufen, wie ein Naturvorgang, noch nach einem vernünftig „verabredeten Plane“ den vorbestimmten Weg gehen. Die Rolle des Geschichtsphilosophen besteht nun

3 *Kritik der reinen Vernunft*, Bd.2, B562. Herausgehoben von mir.

4 Fritz Medicus, „Kants Philosophie der Geschichte“, *Kant-Studien* 7 (1902), 226.

5 *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht* (= *IaG*), Bd.6, A387.

darin, dass er in die un- bzw. widersinnig zu erscheinenden historischen Gegebenheiten „gesetzmäßige Ordnung“ (*IaG*, A392) bringt und ihnen Sinn zu verschaffen sucht.

Kants Idee der Geschichte, die sich in den neun Sätzen der *IaG* präsentiert, soll uns als ein Leitfadens zu einer solchen begreifbaren Geschichte dienen. Vor allem ermöglicht diese Idee, die geschichtlichen Vorgänge so zu betrachten, dass sie nicht als ein blindes Herumtreiben erscheinen, sondern auf einen bestimmten Zweck, nämlich die vollständige Entwicklung aller menschlichen Naturanlagen, abzielen. Als Gattungswesen können die Menschen all ihre „Naturanlagen, die auf den Gebrauch seiner Vernunft abgezielt sind“, entwickeln (*IaG*, A388). Hinter dem ganzen Entwicklungsprozess steckt die Absicht der Natur, die Menschen durch ihren sogenannten „Antagonismus“, „des Lebens und des Wohlbefindens“, also der Glückseligkeit „würdig zu machen“ (*IaG*, A391).

Es gibt nun eine tiefgreifende Meinungsverschiedenheit unter den Kantforschern in Bezug auf das Endziel des geschichtlichen Fortschritts, d. h. ob „die völlige Entwicklung der ‚Anlagen der Menschheit‘“ – durch „die Errichtung einer vollkommenen Staatsverfassung, ... auch die eines weltbürgerlichen Zustandes“ – letzten Endes nur auf „die politische Gerechtigkeit“ gerichtet ist, oder, ob sie „in Moralisierung, d. i. in der Verwandlung des menschlichen Zusammenlebens in ein ‚moralisches Ganzes‘ kulminiert“.<sup>6</sup> Nach den Anhängern der ersteren Interpretationsrichtung soll Kant mit dem Fortschritt der Geschichte ausschließlich die Verbesserung der Institutionen im Auge gehabt haben, insofern die Geschichtsphilosophie mit der Moralphilosophie wenig zu tun haben. War es aber nicht die Ansicht Kants, dass die Menschen die vollkommene bürgerliche Verfassung sowohl auf der nationalen als auch internationalen Ebene errichten sollen, weil sie das „Vehikel der Entwicklung der moralischen Anlagen im Menschengeschlecht“ ist?<sup>7</sup>

Die zweite Interpretationsmöglichkeit wird dagegen mit dem Problem konfrontiert, wie der moralische Fortschritt, der sich *in der Zeit* ereignet, mit den Kantischen moralphilosophischen Auffassungen in Einklang gebracht werden kann. Denn der moralische Fortschritt im-

6 Pauline Kleingeld, *Fortschritt und Vernunft. Zur Geschichtsphilosophie Kants* (Würzburg: Königshausen & Neumann, 1995), 14.

7 Klaus Weyand, „Kants Geschichtsphilosophie. Ihre Entwicklung und ihr Verhältnis zur Aufklärung“, *Kant-Studien: Ergänzungshefte* 85 (Köln: Kölner Universitäts-Verlag, 1963), 106.

pliziert, dass die wesentlichen Elemente der Moralität, wie z.B. das Moralgesetz und das moralische Vermögen des handelnden Subjekts wenigstens gewisse zeitliche Wandlungen zu erfahren haben. Dass der moralische Imperativ von absoluter und ahistorischer Gültigkeit ist, ist aber eine der unverzichtbaren Thesen der Kantischen Ethik. In Ansehung des Menschen als ein moralisches Wesen und seiner praktischen Vermögen, scheint es ebenso wenig plausibel zu sein, von zeitlichen Bedingungen oder Entwicklungen zu sprechen. Wie Kant selber bemerkt, ist es „rätselhaft“, wenn nur die „späteren“ Generationen „das Glück haben ... in dem Gebäude [gemeint ist die vollständige Entwicklung der Anlagen] zu wohnen, woran eine lange Reihe ihrer Vorfahren ... gearbeitet hatten“ (*IaG*, A391). War es wirklich Kants Ansicht, dass eine ausgewählte Gruppe von Menschen, z.B. die späteste Generation der Menschheit, moralisch besser, d. h., freier handeln könne als die anderen?

Das Spannungsverhältnis zwischen „Zeitlosigkeit und Geschichtlichkeit“ wahrzunehmen, heißt nicht unbedingt den Vorwurf erheben zu müssen, Kant sei nicht genügend systematisch. Aufgrund ihrer detaillierten Untersuchung behauptet Pauline Kleingeld, dass gerade in „Kants Konzept der Geschichte als Lernprozeß“ die Vereinbarkeit beider Pole verbürgt ist. Sie meint, „dieses Konzept beinhaltet, daß das Moralitätsprinzip nicht vom historischen Prozeß *geschaffen*, sondern *erhellt* wird.“ Ihr zufolge ist „die Tatsache, daß ein klares Verständnis von Moralität als Autonomie erst im Laufe eines historischen Prozesses erreicht wird, an sich nicht mit dem absoluten Geltungsanspruch des Sittengesetzes unvereinbar.“<sup>8</sup> Die Erhellung bzw. die Verdeutlichung dessen, „was man immer schon Vorbewußt oder dunkel ahnte“, soll aller Vermutung nach die Aufgabe sein, die von der Metaphysik der Sitten durchgeführt wird. Ungeachtet der Tatsache, dass die transzendentalphilosophische Sittenlehre historisch situiert wird, ist demnach der absolute Geltungsanspruch des Moralimperatives von der Fortschrittsthese nicht angetastet.

Wie ist es aber mit dem Moralisierungsprozess des handelnden Subjekts bestellt? Da jede moralisch handelnde Person nach der kritischen Philosophie Kants ein autonomes Wesen, nämlich, ein Noumenon ist, kann eigentlich nicht von einem zeitlichen Fortgang – sowohl auf der Gattungs- als auch auf der individuellen Ebene – die Rede sein. Sollen wir uns vielleicht mit dieser Problematik nicht weiter beschäftigen und dadurch erleichtert fühlen, dass Kant seinen Fortschrittsgedanken bloß als „a

---

8 Kleingeld, 206 f.

regulative idea for heuristic purposes“ präsentiert hat?<sup>9</sup> Man kann also zwar *denken*, dass der Mensch nicht nur als ein Individuum, sondern auch als eine Gattung mit der Zeit moralisch fortschreitet, darf aber nicht die Hoffnung haben, einen solchen Progress *erfahren* zu können? Die Teleologie soll also in die Geschichte nicht deswegen eingeführt worden sein, weil sie die Besserung oder die Steigerung in Bezug auf moralische Gesinnung *in der Zeit* feststellen kann. Sie kann im besten Falle dazu beitragen, dass wir einen bescheidenen (frommen?) Wunsch haben: solche Änderungen sind immer schon im Gang, wenngleich ohne unser Wissen. Wäre das nun alles, was man von dem angenommenen Spannungsverhältnis zu lernen hat?

### 3. Von der „natürlichen Teleologie“ zur „moralischen Teleologie“

Was die Entwicklungsproblematik des Moralsubjekts anlangt, scheint es mir nicht ohne Belang zu sein, sich klarzumachen, welchen Menschen Kant in seinen geschichtsphilosophischen Schriften vor Augen gehabt hat. Im Kontrast zu seinen moralphilosophischen Werken, die hauptsächlich auf die Sicherung der reinen Moralität gerichtet sind, interessiert sich der Geschichtsphilosoph Kant nicht dafür, was der Mensch als ein Noumenon aus sich selbst machen kann und soll. Seine Geschichtsphilosophie macht jedoch nicht ausschließlich den Menschen als Phänomen zum Gegenstand. Ein phänomenologisches Menschenbild dagegen, von dem oben gesprochen worden ist, steht im Vordergrund der geschichtsphilosophischen Überlegung. Das bedeutet, dass Kant hier denjenigen Menschen thematisiert, der sich als ein Doppelwesen darstellt. Der Zweck dieses Interesses beschränkt sich wohl nicht darauf, die Naturanlagen oder die moralischen Werte der Menschen bzw. ihr Verhalten neutral zu beschreiben.<sup>10</sup> Kant begnügt sich z. B. nicht mit dem Hinweis, dass der Mensch nicht nur die Anlage zum Guten, sondern auch die Neigung zum Bösen hat und folglich, wiewohl auf Sinnlichkeit angewiesen, aber dennoch mit der Fähigkeit für freie Wahl begabt, in nuce also ein *animal rationale* ist.

---

9 Pauline Kleingeld, „Kant, History, and the Idea of Moral Development“, *History of Philosophy Quarterly* 16.1 (1999), 70.

10 Das tut Kant m. E. im großen und ganzen in seiner Anthropologievorlesung.

So betrachtet sind Begriffe wie „Moralität“ oder „Freiheit“ usw. in der Geschichtsphilosophie nicht von einer besonderen Art, die sich von der in der kritischen Philosophie unterscheiden lässt. Wenn es irgendeinen Unterschied geben sollte, ist das wohl wahrscheinlich nichts weiter als die Tatsache, dass sie im ersteren Fall in Bezug auf die Menschen behandelt werden, die phänomenologisch begreifbar sind. Die Geschichtsphilosophie will also nicht die der kritischen Philosophie fremdartigen Dogmen liefern, wie z. B. dass der Moralimperativ *in der Zeit* auf eine besondere Weise erfasst wird, oder die Freiheit nicht *ratio essendi* des Moralgesetzes sein kann, usw. Sie will auch nicht im Zusammenhang mit dem moralischen Fortschritt die These vertreten, dass nur die allerletzte Generation der Menschheit autonom sein könne und daher eine moralische Verantwortung und eine Würde von höheren Graden als die anderen habe.

Bei all den Schwankungen und Änderungen kann man eine Beständigkeit in der Botschaft Kants finden, die er durch die Idee des Fortschritts in der Geschichte verkünden wollte. Ursprünglich ging er davon aus, dass die Natur selbst ihren „verborgenen Plan“ bezüglich der Geschichte der Menschheit vollzieht. Das war jedenfalls die Position der *IaG*. Das Mittel, mit dem sie diesen Zweck – die vollkommene bürgerliche Staatsverfassung und dadurch auch die völlige Entfaltung aller menschlichen Anlagen – erreicht, soll die von ihr den Menschen eingepflanzte „ungesellige Geselligkeit“ (Antagonismus) sein. Diese natürliche Teleologie (oder eine Art von Vorsehungsteleologie) tritt zurück, oder verliert wenigstens an Bedeutung, als sich Kant immer mehr der Rolle der frei handelnden Menschen in der Geschichte bewusst wird. Die Geschichte der Menschengattung ist nicht mehr als ein der Natur eigen-tümliches Werk, sondern als das „Werk der Menschen selbst“ anzusehen.<sup>11</sup> Sie kann also nicht mehr die Geschichte der Vorsehung sein. Sie ist die Geschichte der Freiheit der Menschen. Diese Freiheit ist aber nicht von Anfang an den Menschen als ein „Ausstattungsstück“ automatisch zugeteilt. Gegeben ist sie nur „in einem selbst gesetzten Akt des Menschen, durch den er sich von der Verhaftung an die Naturinstinktbindung löst.“<sup>12</sup> Die Natur kann dem Menschen nicht den Endzweck vorschreiben. Eher das Umgekehrte ist der Fall: der Endzweck ist „derjenige

---

11 *Rezension zu Johann Gottfried Herders Ideen zur Philosophie der Geschichte der Menschheit*, Bd.6, A156.

12 Weyand, 135 f.

Zweck, der keines andern als Bedingung seiner Möglichkeit bedarf<sup>13</sup> und daher kann er „nur von einem freien Wesen geschaffen werden.“<sup>14</sup> Ohne Zweifel, nur das moralische Subjekt, das unabhängig von allen Neigungen sich das unbedingte Gesetz aufzuerlegen vermag, kann für ein solches Wesen gehalten werden. Dieser Mensch selbst ist ein Zweck, kann aber zugleich kein Mittel für die anderen Zwecke sein. Mithin ist er, der „seine Bestimmung mit Freiheit ergreift“,<sup>15</sup> der letzte Zweck der Natur. Die Natur erhält ihren letzten Zweck, ihren „Sinn und Wert“ erst durch den freien Menschen, der sich durch die Geschichte kultiviert, zivilisiert, und die Vollendung dieses Prozesses in der Moralisierung findet.

Selbstverständlich sind nicht nur der Naturinstinkt, sondern auch die Anlagen und das Mittel dafür, dass die Menschen von ihrer Fesselung unabhängig sich selbst und all ihre Mitbürger als Zweck an sich betrachten können, von der Natur selbst gegeben. Durch diesen objektiven Tatbestand allein wird jedoch nicht garantiert, dass die Menschengattung bei der Verwirklichung ihrer Bestimmung Erfolg haben wird. Ohne die „moralische Teleologie“<sup>16</sup> wäre aber weder der Sinn der menschlichen Anlagen noch der der Kulturtätigkeiten verstehbar. Diese Teleologie arbeitet in dieser Hinsicht einem geschichtlichen Nihilismus entgegen. Das erklärt aber noch nicht das weitgehende Moment, das Kants Fortschrittsgedanke enthält.

#### 4. Autonomie als Aufgabe: eine „Philosophie des Appells“

Seine geschichtsphilosophischen Überlegungen sind m. E. von der Frage geprägt worden, wie die Aussicht auf die Erreichung der „Bestimmung des Menschen“ durch die Geschichte erhöht werden kann. Im Naturzustand wird ganz vom Zufall her entschieden, mit welchen Naturanlagen und Fähigkeiten und auch in welchem Maße ein Mensch damit begabt ist, und ob er zu ihrer völligen Entwicklung gelangt. Freilich kann auch derjenige, der kein klares Bewusstsein des Moralgesetzes hat, moralisch handeln. In diesem Fall ist aber die Wahrscheinlichkeit gering, dass

---

13 *Kritik der Urteilskraft*, Bd.5, B396.

14 *Medicus*, 193.

15 *Medicus*, 196.

16 *Medicus*, 186.



die Menschen als Gattungswesen irgendwann *in der Zeit* ihre Bestimmung erlangen. Wenn sich die Menschen dagegen in der Geschichte bewusst um die Erhellung und das Verständnis des Moralitätsprinzips bemühen, und die der Idee des geschichtlichen Fortschritts entsprechenden Leistungen vollbringen, werden die Chancen einer Praxis nach diesem Prinzip und somit einer Realisierbarkeit des wirklichen Fortschritts größer.

Angesichts der bisherigen Erwägungen wird klar, wie die Idee des Fortschritts auf Kants Menschenbild wirkt. Kants Geschichtsphilosophie zufolge ist der Mensch kein autonomes Wesen, jedenfalls noch nicht. Sein Freiheitsvermögen ist kein Fertigenmenü, sondern eine zu vollbringende Aufgabe. Von der Dimension der Gattung her gesehen, ist die Autonomie des Menschen die Aufgabe, die er erledigen kann, indem er die äußeren Bedingungen der Moralisierung, nämlich die weltbürgerliche Gesellschaft, Schritt für Schritt errichtet. Ob er die Idee des Fortschritts als ein Postulat anerkennt und dementsprechend handelt, liegt gänzlich in seiner Hand. Wenn diese Auslegung an der Sache nicht vorbeiläuft, wie ich hoffe, dann sollten wir Kants Geschichtsphilosophie, und in gewissem Sinne das ganze System seiner kritischen Philosophie, als eine „Philosophie des Appells“ bezeichnen. Sie appelliert an uns, solche Aufgabe als eine zu realisierende Aufgabe zubeachten.

## 63. Is Kant a Western Philosopher?

*Simon Shengjian Xie*

### 1. Introduction

Some people may find this title nonsensical; they may say, “Germany is a Western country and Kant, being a German, is of course a Western philosopher!” The logic so far is good and I do not blame them. But the truth may not be as simple and clear as the above argument may sound. The reason I pose this question is based on two concerns: the first regarding the definition of the term “Western” and the second regarding the nature or characteristics of Kant’s philosophy in comparison to both Western and Eastern philosophy.

The term “Western” is fundamentally a geographical one, but we also often use it in respect of culture and philosophy that includes a certain “way of living” and “way of thinking”. Here I am mainly concerned with Western philosophy or the Western way of thinking. What is the Western way of thinking? I believe the Western way of thinking can be generalized as the following: in terms of metaphysics, it sees nature or the formation of the world as reducible to scientifically measurable “building blocks” (this being generally known as the “scientific world view”); in terms of epistemology, it sees the origin of knowledge as fundamentally empirical and employs (general) logic as its chief if not the sole method of reasoning whose characteristic is strictly formal, empirical, and analytical. If the Western way of thinking can be defined this way, then Kant cannot be regarded as a Western philosopher because he does not share either of the above two views. Briefly, with regard to metaphysics, Kant is an agnostic, whose fundamental thesis includes the distinction between noumena and phenomena, the former being unknowable; with regard to epistemology, Kant is a transcendental idealist, who employs transcendental logic and a priori synthesis, centering on the a priori relation of knowledge to objects and excluding anything empirical. I will explain more about these two points in the main part of this essay.

The above generalization of the Western way of thinking is obviously best reflected in English-speaking analytic philosophy. Indeed, as English-speaking countries (mainly Anglo-American) are the most powerful (economically and militarily) in the world (with also English being the international language), whose way of thinking as well as way of living has dominated the whole world and when people use the term “Western”, they more or less refer to the dominant Anglo-American. This is why I question the common assumption that Kant is a Western philosopher.

I have briefly mentioned the characteristics of Kant’s philosophy above. To explain it in detail, I will have to discuss the definition of the very framework of philosophy including metaphysics, ethics, epistemology, and logic; in the course of doing so, I will also discuss the main features of Chinese philosophy, a main representative of the Eastern way of thinking as well analytic philosophy. Through this, it will become clear as to what the answer to the title question is and should be.

## 2. Kant’s Philosophical Concerns

Kant’s philosophy is made up of two main parts: his metaphysics, represented by *Critique of Pure Reason*, and his moral theory, represented by *Groundwork for the Metaphysics of Morals* and *Critique of Practical Reason* and some other later works. If philosophy is generally considered as being composed of four areas, that is, metaphysics, ethics, epistemology, and logic, then Kant is mainly concerned with the former two. English-speaking analytic philosophers often refer to *Critique of Pure Reason* as Kant’s epistemology or theory of knowledge, but strictly speaking this is erroneous. This is because, although *Critique of Pure Reason* undoubtedly covers “epistemology”, it is essentially a book of metaphysics; as Kant clearly says in the Preface to the first edition of the book that what he is examining is “the battlefield of these endless controversies” called “metaphysics”.<sup>1</sup> If Kant does have an epistemology, it is practically not the same as what is known to analytic philosophers of empiricist tradition. To clarify these points, we need to look at the terms metaphysics and epistemology.

---

1 Immanuel Kant, *Critique of Pure Reason*, tr. Norman Kemp Smith (London: Macmillan, 1999), Aviii. He adds: “In this enquiry ... I venture to assert that there is not a single metaphysical problem which has not been solved” (Axiii).

## A. The Twofold Meaning of Metaphysics

Metaphysics, as stated at the outset, is one of the four branches of philosophy, and it is more controversial than the other three. The reason is first, since metaphysics deals with ontology, something intangible or transcendental or something “after physics” as Aristotle originally termed it, there is a question about whether or not it exists; second, there is a question about the practicality and validity of the arguments regarding something intangible or transcendental. My answer to the first question is that metaphysics undoubtedly exists and anyone who denies its existence lacks intellectual imagination. My answer to the second question is that I entirely agree with Kant that any arguments concerning metaphysical issues such as God, freedom, and immortality will inevitably fall into antinomies and therefore must be considered to be foolish, useless, and a waste of time. According to this, it is wrong to say that Kant denies metaphysics, as some authors phrase it, because it is far too ambiguous. What Kant denies is not metaphysics itself, but the usefulness and objective validity of the dialectical arguments concerning metaphysical issues.

Philosophers of empirical and analytical tradition tend *either* to deny the existence of metaphysics, because they deny the existence of anything that cannot be scientifically and empirically verified, *or* engage in logical reasoning about metaphysical issues, because they believe that metaphysical issues can be logically justified. These two tendencies in fact come from the same origin, that is, the scientific worldview. This view is so prevalent among contemporary Anglo-American philosophers that they seem to take it for granted. But I will argue that neither Kant nor Chinese philosophy endorse such a view. Such a view is far too simplistic to account for the complexity of our often ineffable experiences of the world. It is high time we took a more sophisticated attitude toward the nature and the makeup of our world and it is high time we went back to Kant and Chinese philosophy for wisdom that is the only possible antidote of the simplistic scientific worldview. In this regard, simplicity here implies stupidity.

## B. Epistemology

Epistemology centers on the question of the origin of knowledge. According to empiricism, experience gives rise to all knowledge; that is, all knowledge must conform to empirical objects—this has become a kind of “folk belief”. What makes Kant distinct is that he famously reverses the role of knowledge and objects; that is, objects must conform to knowledge or concepts. This is known as Kant’s Copernican Revolution. Kant says (B1): “But although all our knowledge begins with experience, it does not follow that it all arises out of experience.” This claim makes room for a priori knowledge—that is, knowledge before experience. What Kant tries to do in *Critique of Pure Reason* is to justify not only the existence of a priori knowledge, but also, more importantly, its objective validity. Kant’s arguments for the objective validity of a priori knowledge are grounded in his metaphysics. At the heart of Kant’s metaphysics lies “noumena” or “things-in-themselves”; they are unknown to us and beyond our access. Because of this, all we can know is within the sphere of phenomena or appearances, corresponding to our spacio-temporal intuitions (i. e., our pure spacio-temporal intuitions become the origin of a priori knowledge). Therefore, Kant famously says (Bxxx): “I have therefore found it necessary to deny *knowledge*, in order to make room for *faith*.” Without faith in noumena, knowledge will go beyond its boundary or go rampant and therefore will no longer be true and valid. If Kant does have an epistemology, as analytic philosophers like to think, then his epistemology is very much subordinate to his metaphysics. Kant’s metaphysical faith in noumena plays the key role in his philosophy, without which his so-called “epistemology” is impossible. This is also what makes Kant a transcendental idealist.

From this we can see that Kant is the enemy of traditional epistemologists of the empirical and analytical schools. If they are the representatives of Western philosophy, particularly in the contemporary sense due to the Anglo-American power, then Kant has nothing in common with it.

### 3. Epistemology as Surrogate Metaphysics

The greatest achievements English-speaking people have made to the world (apart from English language and literature) are in science and technology; the comfort and convenience they provide to people's daily lives can be seen by all. It is fair to say that English-speaking people are scientifically and logically minded. But philosophy is more of an intellectual and abstract domain that is not directly linked to people's daily lives, or at least its effect on us is not immediate but far-reaching. But scientifically and logically minded people tend to be short-sighted; they want to see immediate effects in everything, even in philosophy. So they try to "scientifically" philosophy, turning it into a scientific enterprise. This explains why Western philosophy (analytic philosophy in particular) is strong in epistemology: it treats epistemology as a science and aims at logically justifying epistemological arguments. Since metaphysics cannot be logically justified and cannot be treated like science, analytic philosophy betrays its weakness by doing metaphysics the same way it does epistemology and science. If metaphysics is the essence of philosophy, then analytic philosophy lacks essence. In fact, analytic philosophy tends to use epistemology to replace metaphysics, thus reducing the unknowable or the "ineffable" in metaphysics to the scientific worldview. While I acknowledge that analytic philosophy is strong in epistemology and logic, I would also like to point out that it is weak in metaphysics and ethics because of its non-believing and impractical approach. Metaphysics cannot be replaced by epistemology and neither can moral axioms and principles be replaced by logical premises, propositions, and reasoning. Any efforts toward this effect practically destroy metaphysics and ethics—and philosophy for that matter. But sadly, this is what I believe certain analytic philosophers are doing.

### 4. The Significance of the Parity

I have hinted in my above argument that logic and epistemology go hand in hand and ethics and metaphysics go hand in hand. This is indeed what I would like to point out. Briefly, epistemology can be argued, and logic is the reasoning tool that must therefore be employed by epistemology. Metaphysics entails belief or faith that defies arguments, so it does not need logic; but it can be practiced in human action. Ethics is about human action or conduct in the human world, so it must be

paired with metaphysics as its empirical outlet. This pairing is significant because it sets the task or function of each of these philosophical branches, and therefore also its limit.

Analytic philosophers rarely talk about these pairings, and as a result, they often take it for granted that logic is to be used as a universal tool for all philosophical areas, including metaphysics and ethics. In doing so, they create new fields as ridiculous as meta-metaphysics and meta-ethics. Although meta-ethics is a more established area than meta-metaphysics, this does not make it legitimate. The reason is that moral right and wrong at the fundamental level comes from the moral sense of rational beings that is a priori and is related to their metaphysical view of the world: you either have it and have faith in it or do not have it and have no faith in it. This is why there is such a thing as a moral law, functioning as an “axiom”. Just like mathematical acumen, a person either has it or does not have it, and there is hardly any room for argument. Why does one have to believe that  $1 + 1 = 2$ , or  $1 + 1 = 10$  as in the binary system? Those who do not have mathematical acumen do not mess around in mathematics because its boundary is clear and no one wants to make a fool of oneself. But those who do not have or do not believe in moral sense of good and bad or right and wrong do mess around in the moral domain, due to its apparently not-so-clear boundary. Thus there arise those so-called meta-ethicists. To put it another way, meta-ethicists are simply logicians messing around with fundamental moral issues.

Of course there is also an area called “applied ethics” where many ethical issues can be argued, such as abortion, euthanasia, and various environmental ethical issues. These are practical issues but not fundamental issues; therefore they can be treated more flexibly through debates (different societies may even have different attitudes and laws towards these issues), whereas the fundamental issues such as the moral law, or good will as the only good in the universe without qualification, cannot be so treated. Doubting the validity of such issues simply proves that a person does not know what moral good is and does not understand the nature of morality in the first place. Just like when a person doubts the existence of God, this simply shows that the person is non-religious in the first place. This also explains why the majority of analytic philosophers are skeptics. Skepticism can be good in challenging unfounded or dark-age kinds of heresy or dogmatism. But it can be extremely bad in not having any faith at all and in not accepting beliefs originated from deep insights and profound wisdom. One may ask:

how can we tell the latter from the former? Well, having insight, wisdom, or intellectual intuition is the answer. Skepticism is the easy option for the intellectually or rationally weak-minded.

## 5. Logic and Philosophical Methods

Now let us come back to logic. Although logic is also supposed to be an intellectual and abstract branch of learning, its foundation is essentially psychological or emotional because it relies on “conditions” to proceed and it cannot do without procedures. Therefore its nature is still “concrete” or something that can be measured or justified. Logical reasoning is thus the main weapon of analytic philosophy. But logical reasoning is not the only method in doing philosophy. Analogy and metaphor among others are just as important, if not more important. Many philosophers regard poetic language as the most suitable language for philosophy. In fact, good analogy and metaphor come from intellectual imagination, insights, and wisdom. It is little wonder that analytic philosophers deny analogy and metaphor as valid philosophical methods because they love logical reasoning more than wisdom, as one renowned analytic philosopher admits.<sup>2</sup> But love of reasoning cannot replace love of wisdom, for the latter is the true definition of philosophy and goes hand in hand with deep insights and creative thinking; these latter we may also regard in general as intellectual intuition. (This, in fact, can settle the issue whether the Chinese have philosophy. When some philosophers say that the Chinese have no philosophy, they are clearly prejudiced and have misconceived what philosophy is. They obviously take it for granted that philosophy means love of reasoning instead of love of wisdom. They would have to be ashamed of themselves if they realized how much wisdom is involved in Chinese ways of thinking and how close it is to the true meaning of philosophy.) True philosophy must involve intellectual intuition, not just logical reasoning and analysis. Without the former, philosophy will become shallow, purposeless, and soulless.

---

2 Robert Nozick, *The Nature of Rationality* (Princeton: Princeton University Press, 1993), xi, says: “The WORD *philosophy* means the love of wisdom, but what philosophers really love is reasoning.”



## 6. Intellectual Intuition and Chinese Philosophy

By comparison, the main concerns of Chinese philosophy are metaphysics and ethics, just like those of Kant. To be more exact, Daoism is the metaphysics of Chinese philosophy, or its cosmological view of the world, and Confucianism is the ethics of Chinese philosophy, regulating individual conduct in a reality governed by transcendental ideas. This is to say that, as far as his main philosophical focus is concerned, Kant exhibits eastern or Chinese characteristics. Daoism is the metaphysics of Chinese philosophy and is renowned for its illogical, metaphoric, and analogous way of thinking or reasoning. If one has little idea as to what metaphysics is like, one only needs to think of the picture *Dao De Jing* tries to draw. The Dao of *Dao De Jing* is, in fact, what Kant calls noumena; it is unknowable but can be thought, according to Kant, but it is understandable to Lao Zi. In this sense, Daoism starts from the point where Kant's philosophy ends.

The reason Kant thinks we cannot know noumena or things-in-themselves is that we do not have intellectual intuition. Here Kant underestimates human intellectual power. In fact, Chinese philosophy relies on intellectual intuition. Kant is not wrong in saying things-in-themselves are unknowable, because to know things-in-themselves, we have to be able to explain the concept in plain language, but this is indeed beyond us. Lao Zi does not pretend to "know" Dao or things-in-themselves as such. Otherwise *Dao De Jing* will be an easy read. What intellectual intuition can do is not to know Dao in plain language, but to imagine what it is like and to understand its working so that people can follow its law and utilize it for humanity's own good. To convey one's imagination is nothing like speaking one's thoughts. Therefore, the language of *Dao De Jing* is full of metaphors, analogies and elusive language. In short, *Dao De Jing* is the product of the imagination of intellectual intuition. As far as his denial of our having intellectual intuition is concerned, Kant is still a Western philosopher. But as far as his recognition of the existence of noumena or things-in-themselves is concerned, Kant has become an Eastern philosopher. (Here I need to point out that, although Kant denies our having intellectual intuition, he himself has already shown sparks of it. Otherwise, transcendental idealism would be impossible. By contrast, in *Dao De Jing*, Lao Zi's intellectual intuition is in full bloom.)

In conclusion, the reason I argue for the claim that Kant is more of an Eastern philosopher should now be clear. On the whole I have found

that Anglo-American analytic philosophy has usurped the role of true philosophy when it is only a school of philosophy, one that may be more appropriately called “applied philosophy”. I hope to have shown that in the overall framework of philosophy, both Chinese and Kant’s philosophy occupy the more authentic parts of philosophy, or philosophy proper. By revealing this truth, I hope to have provided philosophers with a warning not to be blinded by the appearance of the apparent dominance of Anglo-American analytic philosophy, but to have faith in the wisdom and insights of Kant and Chinese philosophy and to show more interests in them in order to save philosophy from being downgraded to mere logical and linguistic experiments.

### **The Hierarchical Structure of Philosophy**

*Metaphysics proper*: the “ineffable” or noumena

*Epistemology*: dialectic reasoning

*Morality proper*: the moral law as universal

*Applied ethics*: empirical and socio-cultural

*Transcendental logic*: categorical and systematic

*General logic*: conditional and procedural



## EPILOGUE



## 64. The Unity of Architectonic Reasoning in Kant and *I Ching*

Stephen R. Palmquist

Human reason is by nature architectonic, i. e., it considers all cognitions as belonging to a possible system, and hence it permits only such principles as at least do not render an intended cognition incapable of standing together with others in some system or other.<sup>1</sup>

The unity of human personhood in Kant's philosophical system is not incompatible with a belief in the duality of human nature, nor with an appreciation of the fragmented nature of our empirical existence. He portrays human beings as belonging simultaneously to both the phenomenal and the noumenal "worlds", as being causally determined by events in the *natural* world that we cannot control, yet having the spontaneous power to initiate freely chosen actions that constitute a *moral* world. Likewise, he makes numerous finer distinctions between various types or aspects of human character or personality throughout the three *Critiques*, as well as in his minor writings, lectures, and notes. As the foregoing essays demonstrate, we find in each *Critique* and throughout Kant's writings a sometimes mesmerizing array of distinctions regarding our nature as human beings, yet each *Critique* is united by its focus on one of three central questions (A805/B833): "1. What can I know? 2. What should I do? 3. What may I hope?" Moreover, these questions are united by one question that combines them all: "What is man?"

---

1 A474/B502. For a thoroughgoing discussion of Kant's use of the term "architectonic" and of the nature and structure of the architectonic plan he used to structure the Critical System, see my books, *Kant's System of Perspectives: An architectonic interpretation of the Critical philosophy* (Lanham: University Press of America, 1993), ch. III, and *Kant's Critical Religion: Volume Two of Kant's System of Perspectives* (Aldershot: Ashgate, 2000), apx.III.2–3.

(9:25)—a deceptively simple question that seems to call for one, all-encompassing answer.

Kant's clue as to how we can, paradoxically, both believe in the fundamental *unity* of human personhood and acknowledge the seemingly endless "aggregate" of unorganized facts that characterizes our human nature is that Critical philosophers must employ a special kind of thinking he calls "architectonic". My purpose in this essay is not to describe how architectonic thinking manifests itself in all Kant's intricate theories of human personhood—that task has already been fulfilled by the many insightful essays contained in this collection. Rather, after discussing Kant's special, architectonic approach to philosophical reasoning and its systematic relation to the twelve categories, I shall suggest that the same approach can be found, in its essential nature though not in its detailed out-working, in the oldest and arguably the most influential of all Chinese classics, the *I Ching*. If I am correct, then the "Chinaman of Konigsberg"<sup>2</sup> was even more authentically Chinese than either he or Nietzsche realized.

The *I Ching* or *Book of Changes*, as it is sometimes called, is essentially a set of commentaries on sixty-four unique hexagrams that are each constructed out of six broken or solid lines. Because each line can appear in only two forms (broken or solid) and each component of the system contains exactly six such lines, the sixty-four hexagrams represent *all possible* permutations of any system exhibiting this logical form. This is true for the simple, mathematical reason that  $2 \times 2 \times 2 \times 2 \times 2 \times 2$  (i. e.,  $2^6$ ) is 64. The Chinese classic interprets each hexagram as a symbol representing a certain human situation or type of situation, based on its unique arrangement of broken and solid lines. Those who use the *I Ching* as a guideline for decision-making (or, more crassly, as a tool for divining the future) adopt this set of 64 logical relations as an *a priori* framework by randomly choosing two of the 64 hexagrams and viewing them as a symbolic representation of the *change* being exhibited by some situation they wish to understand more fully. Mastering the *I Ching* requires one to learn the nuances of 4,096 (i. e.,  $64 \times 64$ ) mathematically possible types of situational change generated by the logic of this system. Although I am still a novice in this regard, I shall illustrate at the end of this essay how such applications might operate as a practical, architectonic guideline.

---

2 I assess the legitimacy of viewing Kant as taking a "Chinese" approach in "How 'Chinese' Was Kant?" (abridged), *The Philosopher* 84:1 (Spring 1996), 3–9.

An intriguing feature of this ancient framework for interpreting human experience is that it all arises from a fundamental *unity*, called the Dao (though the Dao paradoxically also underlies the very distinction between unity and diversity). As expressed in the well-known lines from section 42 of Lao Tzu's *Dao De Jing*:<sup>3</sup>

DAO generates the One  
 The One generates the Two  
 The Two generates the Three  
 The Three generates all things.

Reading these lines in connection with their roots in the *I Ching* provides an excellent expression of the book's underlying assumption, that unity and diversity are not necessarily incompatible concepts, but can work together to elucidate how we experience human life as a coherent whole.

This well-known passage has some interesting implications for the question of how the unity of human personhood can coexist with the diversity of life as we experience it; but what has any of this to do with Kant? It has to do with Kant because he famously (or by some accounts, infamously) insists that philosophers ought *not* interpret the world in the manner of Aristotle, by merely collecting data from our observations of the world and inductively classifying these according to some likely set of conclusions, but should rather *impose order* onto our subject-matter through a predetermined principle of division that *we give* to the system of concepts we employ.<sup>4</sup> The *Dao* on its own is a name for undifferentiated wholeness, not unlike Kant's "thing in itself". We come to *know* it as "one", "two", "three", and eventually "all things", only by imposing our mental categorizations (the 64-hexa-

---

3 Lao Tzu, *Tao Te Ching: The Book of Meaning and Life*, tr. Richard Wilhelm and H. G. Ostwald (London: Arkana, 1985), 46.

4 For a good account of the tendency among early twentieth century Kant scholars to blame all the infelicities one sees in Kant's writings on his architectonic superstructure, see Paula Manchester, "What Kant Means by Architectonic" *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses*, Bd. II (Berlin: Walter de Gruyter, 2001), 622–30. Manchester's own interpretation of Kant's view of architectonic is, however, unfortunately clouded by her overly Aristotelian reading of Kant's usage, combined with an overemphasis on the significance of Kant's reference to the "teaching" of reason in connection with architectonic. As I shall demonstrate, Kant explicitly *contrasts* his position with Aristotle's "aggregate" approach, so it seems highly unlikely that he saw himself as merely refining the same meaning Aristotle gave to this term.



grams, in classical Chinese traditions) onto it. This—dare I say?—*Kantian* aspect of philosophical Daoism might go unnoticed if we interpret it apart from its relation to the *I Ching*. Likewise, Kant’s unified answer to the “What is man?” question is likely to remain obscure if we do not recognize how his table of twelve categories originates as a presupposition of architectonic reasoning. Let us therefore look first at chapter III of *CPR*’s Transcendental Doctrine of Method, where Kant explains what he means by “architectonic”, then examine how he provides early hints concerning this important philosophical method in the untitled introductory section of chapter I of the Transcendental Analytic, where he explicitly refers to the “clue” that leads to the discovery of the categories.

In chapter III of *Kant’s System of Perspectives*<sup>5</sup> I have presented in great detail the multi-layered structure of Kant’s so-called “architectonic plan” for constructing his philosophical system, arguing throughout the rest of the book that commentators who misunderstand and prematurely reject Kant’s theories typically do so because they fail to appreciate how his various arguments contribute to this plan as a whole. After being criticized by Paula Manchester for misunderstanding Kant’s use of the term “architectonic”,<sup>6</sup> I wrote a detailed reply, in appendix III.2 of *Kant’s Critical Religion*,<sup>7</sup> presenting a more explicit account of what Kant means by this key term. Without repeating all the details of those two studies, I shall here summarize the key features of Kant’s position.

Kant begins the Architectonic chapter with this intriguing definition: “By an **architectonic** I understand the art of systems. Since systematic unity is that which first makes ordinary cognition into science, i. e., makes a system out of a mere aggregate of it, architectonic is the

---

5 For the full text of *Kant’s System*, see <http://www.hkbu.edu.hk/~ppp/ksp1>.

6 Manchester views Kant’s “architectonic” as essentially following Aristotle’s use of the special Greek term, *architektôn* (Manchester, 524n). After reviewing the history of different uses of this term, she assumes Kant must have been writing within this Greek tradition; unfortunately, she never presents such a thorough analysis of the distinctive way *Kant himself* uses the term, especially in light of his explicit contrast between his view of how to construct a table of categories and that of Aristotle (see note 4, above). Manchester and I debated this issue at a special session of the 1998 World Congress of Philosophy, but without reaching agreement.

7 For the full text of *Kant’s Critical Religion*, see <http://www.hkbu.edu.hk/~ppp/ksp2>.

doctrine of that which is scientific in our cognition in general” (A832/B860). Here we see Kant connecting architectonic with system-making, apparently hinting at a metaphor between the roles of the architect and the philosopher. Just as an architect’s job is to design or “make the plan” for a building, the architectonic philosopher’s task is to make *systems* by imposing order onto the “mere aggregate” (i. e., the unorganized data) that otherwise characterizes our experience of the empirical world. Perhaps the most intriguing aspect suggested by this passage is that Kant calls architectonic an “art”, even though it is at the same time, somewhat paradoxically, the formal factor that makes a body of knowledge *scientific*. He then adds that reason *prescribes laws* that unify “the manifold cognitions under one idea” (A832/B860). This idea, he tells us, is “the rational concept of the form of a whole” that determines both “the domain of the manifold” and “the positions of the parts with respect to each other”. That is, the task of architectonic reasoning is to *determine the relation* between the otherwise unrelated parts of a transcendental system’s form.

Two sentences later Kant again emphasizes this relational aspect. Apparently, he had an architectonic reason for placing this chapter *third* in the Doctrine of Method: it fulfills a function that corresponds to the category of relation in his Table of Categories. As I argue in chapter VII of *Kant’s System of Perspectives*, the component of the Doctrine of Elements that functions as the architectonic structuring plan is the categories, applied in the schematized form of the principles of pure understanding. Here in the Doctrine of Method, Kant therefore appears to be alluding to a necessary connection between the formal structure of the categories and that of all architectonic reasoning. If this interpretation is accurate, then why did Kant not simply come out and state that architectonic reasoning uses the table of categories (or its predecessor, the table of the logical forms of judgment in thought) to impose systematic patterns onto our thought processes? The reason, I believe, is bound up with Kant’s strategy in dividing the *Critiques* into Doctrines of Elements and Method. In each *Critique* with this division, the two sections are meant to be *independent* of each other, in the sense that they work toward the same goal, but from opposite perspectives: content first, then form. None of the chapters in the Doctrine of Method appeal directly to the results of the Doctrine of Elements, or vice versa; rather, they each reveal in different ways reason’s *need* for just the sort of thing the foregoing Doctrine of Elements has provided. To connect architectonic in chapter III of the Doctrine of Method too explicitly with the

$4 \times 3 = 12$  pattern determined by the categories in the Doctrine of Elements would have been to beg the question he was attempting to answer. To *name* the categories or even their numerical structure would have been to focus on the *content* of the preferred architectonic plan; but Kant's focus in the Doctrine of Method is on the proper *form* of philosophical reasoning, a form that could be different for different philosophers.

The second paragraph of the Architectonic also states that the *purpose* of imposing onto the aggregate of our knowledge an idea that relates the parts to each other within a whole is to “support and advance [reason’s] essential ends” (A832/B860). Kant unfortunately does not explain what he means by this phrase. However, the remainder of the paragraph suggests he is thinking here of reason’s ultimate goal, the *unification* of all knowledge; for he claims this prescriptive function of reason (i. e., reason’s architectonic unity) “allows the absence of any part to be noticed in our knowledge of the rest,” so that “there can be no contingent addition ... that does not have its boundaries determined *a priori*” (A832–3/B860–1), thus guaranteeing the completeness of the system being constructed. In the Doctrine of Elements, the only tool Kant develops for achieving such lofty aims is his choice to pattern his systematic divisions on the formal structure established by the tables of categories and logical functions. Perhaps hinting at his own earlier usage, he concludes this paragraph of the Doctrine of Method by comparing a rational system’s potential to “grow internally ... but not externally” (i. e., to be “articulated” rather than “heaped together”) to that of “an animal body” (A833/B861). This metaphor is easily understood as referring to Kant’s conviction that, when constructing a table of categories in reference to any set of conceptual relations, we must resist the temptation to add a single new member (e. g.,  $4 + 1 = 5$ ), for this destroys the logical unity of the conceptual relations under consideration. Instead, we must account for any new members by making further *internal* divisions, just as Kant does when he divides each category into three “moments” ( $4 \times 3 = 12$ ).

The third paragraph contains the next two references to “architectonic”. It begins by distinguishing between two ways of relating a *schema* and an *idea*. Viewed empirically, the schema presents the manifold of knowledge to us independently of any unifying idea, whereas from reason’s *a priori* perspective, the schema “arises only in consequence of an idea ... and does not await them empirically” (A833/B861). The latter alone, Kant states, “grounds **architectonic** unity.” One of the main

differences between these two forms of relation is that when viewing the schema “empirically”, we cannot know the “number [of its aims] ... in advance”; but science requires certainty in its distinctions and so must impose them *a priori*—that is, “architectonically, for the sake of its affinity and its derivation from a single supreme and inner end” (A833–4/B861–2). This passage provides clear evidence that the *a priori* unity imposed on the aggregate by reason’s architectonic art has to do with the  $4 \times 3 = 12$  pattern of the categories. For Kant’s point is precisely that reason’s architectonic form (as revealed in the categories) enables us to do what would be impossible if we were to use a merely empirical method: to *determine the appropriate number* that composes any given set of concepts. Reason’s ability to discern the pattern in advance is the source of the affinity of the manifold’s parts in an architectonic system.

The fourth paragraph warns the reader that, although the founder of every new science bases it on an idea, the initial attempt to *schematize* that idea “seldom corresponds to the idea; for this idea lies in reason like a seed” (A834/B862). As a result, Kant encourages us to be willing to go beyond the descriptions given by the founders and first proponents of any new science, for they “often fumble around with an idea that they have not even made distinct to themselves”; our focus should instead be on the idea and its grounding in reason. This accords well with my articulation of the logical structure of the architectonic form of Kant’s System, given in chapter III of *Kant’s System of Perspectives*. If Kant is to avoid being hypocritical, he would have to confess that he, too, like the founder of any new science, had only a vague grasp of the “idea of the whole” that brought unity and completeness to his System of transcendental philosophy. (My goal in *Kant’s System of Perspectives* was to apply this advice of Kant’s to the task of interpreting the architectonic structure of his own System.)

The next three occurrences of “architectonic(ally)”, coming in the fifth paragraph of chapter III, do not tell us anything fundamentally new about Kant’s understanding of the term. The paragraph begins by lamenting that systems are typically constructed initially as aggregates, and that “only after we have long collected relevant cognitions haphazardly like building materials” does it first become “possible for us to glimpse the idea in a clearer light and to outline a whole architectonically, in accordance with the ends of reason” (A834–5/B862–3). (The fact that Kant made essentially the same point in the so-called Metaphysical Deduction [A79–81/B105–7], in criticizing Aristotle’s meth-

od of collecting categories, provides yet further evidence that the table of categories is Kant's most complete expression of the formal structure he prefers when employing architectonic reasoning.) After likening the development of aggregate systems to the work of "maggots" (A835/B863), he claims that "so much material has already been collected" in relation to human cognition that giving "an architectonic to all human knowledge ... would not only be possible but would not even be very difficult." He then says the remainder of the chapter will merely outline "the **architectonic** of all cognition from **pure reason**".

Without looking any further into the details of the Doctrine of Method's Architectonic chapter,<sup>8</sup> let us now turn to the Doctrine of Elements, to chapter I of the Transcendental Analytic, where Kant first introduces his perplexing "clue to the discovery of all pure concepts of the understanding" (A66/B91). In the short (untitled) introductory section, Kant explicitly compares the "mechanical" method of attempting to find completeness among the manifold concepts that arise out of our observations of the world with the special method adopted by the transcendental philosopher. The former method is the one Aristotle used to gather and present his list of categories; such "concepts that are discovered only as the opportunity arises will not reveal any order and systematic unity, but will rather be ordered in pairs only according

---

8 From this point much of chapter III consists of a series of twofold divisions of reason and/or philosophy, intended to provide the reader with a bird's eye view of the architectonic form of transcendental philosophy. We can skip over the details of Kant's exposition, not only because the various divisions appear at times to be somewhat incompatible with each other, but also because they are advanced as *examples* of architectonic divisions, not as further explications of the meaning of the term as such. Instead of recounting the details of each division, we can pass on to Kant's final use of "architectonic" in *CPR*. Six paragraphs before the end of chapter III, immediately after summarizing "the entire system of metaphysics" in terms of "four main parts" (A846/B874), Kant reaffirms several aspects of the meaning of "architectonic" I am defending here (A847/B875): "The original idea of a philosophy of pure reason itself prescribes this division; it is therefore **architectonic**, in conformity with its essential ends ...; and for that very reason [this division] is unchangeable and legislative." Once again we see that this term entails that reason has *prescribed a division* (i. e.,  $4 = 2 \div 2$ ) "in conformity with its essential ends"; because it conforms to reason's ends (i. e., to the categories as applied in the principles), this division can be regarded as authoritative and "unchangeable". Kant also uses the term "architectonic" in a number of his other writings; for a further discussion of these references, see Appendix III.3 of *Kant's Critical Religion*.

to similarities . . . , from the simple to the more composite” (A66–7/B91–2). By contrast, the latter “has the advantage but also the obligation to seek its concepts in accordance with a principle”; Kant thus adopts this approach to produce a table of categories consisting of concepts that “spring pure and unmixed, out of the understanding, *as absolute unity*” (A67/B92, emphasis added). The resulting table illustrates the correct procedure for architectonic philosophizing, while that procedure constitutes the “clue” to understanding why Kant thinks the twelvefold table of categories is complete in the *form* he presents it. Adopting such a predetermined, architectonic plan is the only way to avoid a situation where the choice of basic concepts depends merely “upon whim or chance.”<sup>9</sup>

I shall now conclude with some further reflections on the *I Ching*, based on an experimental application of the latter to the thesis advanced in this essay, that the unity of the *I Ching* is also based on a predetermined, *a priori* form. At the risk of appearing foolish to any interpreters who are not yet convinced that one *must* take into account Kant’s belief in the architectonic nature of correct philosophical reasoning, if we are to interpret his philosophical doctrines accurately, I shall treat the *I Ching* as itself offering us an architectonic plan (though its form is clearly different, based on 2<sup>6</sup> rather than on 3x4) and will “ask” it a specific question about the unity of architectonic reasoning. By randomly selecting two hexagrams in the manner mentioned at the beginning of this essay, I hope to shed further light (or cast further doubt) on the usefulness of architectonic reasoning.

An interesting characteristic about the *I Ching* is that it *appears* to be based on chance. For example, at 3am on the night before the Kant in Asia conference began, I used sixteen colored marbles to select one of the 64 hexagrams, while thinking about the following question:

---

9 A67/B92. These two methods are aptly illustrated by an example Kant provides in the Second Preface of *Religion within the Bounds of Bare Reason*, where he says reason and Scripture should have “not only compatibility but also unity” (6:13, tr. Pluhar, 2009). For an overview of how Kant’s architectonic pattern applies to the text of this book, see my “Introduction” to Pluhar’s translation (Indianapolis: Hackett, 2009). Exactly how this “unity” arises (or might arise) is a question Kant does not clearly answer in that context. But if we understand the way architectonic reasoning operates, the problem of the unity of “the religion of reason” with empirical religious ideas can be easily solved. See Love’s essay (ch. 42, above) for a discussion of how this relates to the problem of religious pluralism.

“What is the likely result of an attempt to connect Kant’s theory of the unity of architectonic reasoning, as manifested in his table of categories, with the formal structure of the *I Ching*?” The immediate result of my little experiment was, indeed, random in the sense that I could have ended up selecting any one of the 64 possible combinations of marbles. One might argue that this is so different from Kantian architectonic as to be totally irrelevant. But wait. Kantian categories do not remove the randomness and contingencies of our day-to-day experience; they only help us understand how the diversity of empirical knowledge can be unified. Should we not give the *I Ching* an equal chance?

My choice of marbles ended up presenting me with hexagram 21, changing into hexagram 38. Number 21 is called “biting through”; it shows an open mouth with an obstruction. The maxim for this hexagram reads: “Energetic biting through overcomes the obstacle that prevents joining of the lips.”<sup>10</sup> This suggests that the attempt to reconcile the opposing points of view (of Kant and the *I Ching*—and ultimately, of Kant and Asian philosophy) therefore seems possible, but will require hard work. This first hexagram represents the situation I, the asker, had come from: during the several months prior to the conference I had found the need, as Convener, to “bite through” several obstacles. Significantly, the second hexagram (number 38) is called “Opposition”. While this may appear to be not very auspicious, we should not make such an assumption too hastily. At one level, it seems almost as if the message conveyed by this hexagram ended up predicting the future: after the conference, a colleague whose preferences I had “bitten through” opposed me so strongly that he lodged a formal complaint against me. However, the question I asked the *I Ching* was not personal; so let us instead consider this deeply Kantian message that happens to be conveyed by the commentary on hexagram 38:

In general, opposition appears as an obstruction, but when it represents polarity within a comprehensive whole, it has also its useful and important functions. The oppositions of heaven and earth, spirit and nature, man and woman, when reconciled, bring about the creation and reproduction of life. In the world of visible things, the principle of opposites makes possible the differentiation by categories through which order is brought into the world.<sup>11</sup>

---

10 *I Ching or Book of Changes*, the Richard Wilhelm translation, tr. Cary F. Baynes (London: Routledge and Kegan Paul, 1951), 148.

11 *I Ching*, 148.

Just how different is this use of architectonic reasoning from that adopted by Kant? Clearly, they are not the same. But we should not expect them to be identical, given that the *I Ching* predates Kant by several thousand years. Kant's employment of the categories served as a transcendental basis for understanding the modern scientific and religious worldview (see note 9, above). The *I Ching* did (and does) nothing of the kind, for empirical science was (at best) in its infancy when the system of 64 hexagrams was first conceived. Nevertheless, it does exemplify a *method of thinking* that is remarkably similar to Kant's. Kant's predetermined divisions in philosophy (especially the categories) *lead us* into insights about science and religion, just as the random selection of hexagrams, when interpreted as a predetermined set of symbols describing 64x64 life situations, can often lead us into remarkable insights about how to understand any given life situation.

The paradox we face when attempting to employ architectonic reasoning also constitutes what is arguably the single most dangerous temptation faced by philosophers (or by anyone thinking philosophically). We always have the tendency to believe that our structured understanding of the nature of reality (or of any given situation) represents the absolute truth. It is no accident, perhaps, that the *I Ching's* reputation has been spoiled in so many circles: the hexagrams are often used explicitly for divination, as if we human beings could know the future simply by casting yarrow sticks (or grabbing marbles out of a bag). Yet if we resist this temptation, employing architectonic reasoning without forgetting that we have created the structures in the first place, then it can be the source of great wisdom and insight. In such uses, we actually are "divining the truth" by imposing an architectonic structure onto the empirical aggregate. Without adopting this approach, we can never hope to find unity in the midst of our diverse efforts to cultivate personhood. Yet the lesson of Kant's Critical philosophy is that (as aptly expressed by my friend Guy Lown, one of the participants in the Kant in Asia conference, in a discussion we had on this topic just as I was finalizing this collection of essays) even though the purpose of architectonic systems is to *divine the structure of reality*, we must learn to do this *without* regarding the outcome of our reasoning as divine. I can think of no better way of realizing this goal than by observing (architectonic reasoning being but one of many examples of) how Kant's ideas are alive in Asia and Asian ideas resonate in Kant.



## Note on Contributors

This collection consists of essays by three keynote speakers, followed by revised versions of submitted essays by 64 regular participants. They are listed here in alphabetical order.

### Three Keynote Scholars

**Chung-ying Cheng** received his PhD in Philosophy from Harvard University in 1964. He has been Professor of Philosophy at University of Hawaii at Manoa since 1972. He founded the *Journal of Chinese Philosophy* and has served as its editor-in-chief since 1973. He also founded the International Society for Chinese Philosophy. Professor Cheng's recent works cover onto-hermeneutics, Yijing, Daoism, and Confucianism/Neo-Confucianism. He has published 25 books in both English and Chinese and more than 300 papers in various fields of philosophy, including Contemporary Chinese Philosophy, Theory of Confucian Philosophy, Creating Harmony, Ontology and Interpretation (6 volumes), and Philosophy of Yijing Ontology.

**Patricia Kitcher** is Mark van Doren Professor of Humanities and Professor of Philosophy at Columbia University. After receiving her PhD from Princeton in 1974, she taught at the University of Vermont, the University of Minnesota and the University of California, San Diego. From 2002–2005, she served as President of the North American Kant Society. She is the author of numerous books and articles, including *Kant's Transcendental Psychology*, *Freud's Dream: A Complete Interdisciplinary Science of Mind*, and most recently *Kant's Thinker*. Her current project is on Kant's moral agent.

**Günter Wohlfart** studied philosophy, German literature, and psychoanalysis in Germany. While teaching at various German universities, he published books on Kant, Hegel, Nietzsche, and Heraclitus. After his transcultural turn he focused on Zen Buddhism and philosophical Daoism, publishing books on Zen, Laozi, and Zhuangzi. As a guest professor he taught in China, Japan, Korea, and Taiwan. Since retiring from his latest position, as Professor of Philosophy at Wuppertal University, Wohlfart has lived in his mountain cottage in southern France. He

spends his time travelling, writing, and gardening. His latest research is an intercultural study concerning a polemic against Kant's ethics.

### Sixty-Four Contributors

**Claudia Bickmann**, Professor of Philosophy at Cologne University, Germany, has been Visiting Professor in Prague, Delhi, Cairo, and Alexandria. Her books deal with Literary Theory, Kant's Philosophical System, and Kant's World-Philosophy. Her numerous articles are on Intercultural Philosophy, Plato, Aristotle, Kant, Fichte, Schelling, Hegel, and Heidegger. She is President of the International Society of Intercultural Philosophy.

**Alexander Bobko** is Professor of Philosophy at University of Rzeszow and at The Pope University of John Paul II in Krakow, and Vice-Rector of University of Rzeszow. His 50+ publications, mostly on Kant, include five books, a Polish translation of *Religion innerhalb der Grenzen der blossen Vernunft*, and multiple articles published in Germany, USA, and elsewhere.

**Annie Boisclair** is a PhD student in philosophy at the University of Montreal. Her research addresses Mou Zongsan's use of Perfect Teaching. She studied "good and evil" in *Laozi* and *Zhuangzi* as a Master student at Taiwan's National Central University. She has published an analysis of symbols illustrating the Autumn Floods of *Zhuangzi*.

**Andrew Brook** (DPhil, Oxford) is Chancellor's Professor of Philosophy and Cognitive Science at Carleton University, Ottawa. He is former President of the Canadian Philosophical Association and former Director of the Institute of Cognitive Science at Carleton. Among his 100+ publications are *Kant and the Mind* and about 15 articles on Kant. Website: [www.carleton.ca/~abrook](http://www.carleton.ca/~abrook).

**Monique Castillo** is Professor of Philosophy at University of Paris-East and graduate of the Institute of Political Sciences (Paris). She has written: *Kant and the Future of the Culture*; *Kant, the Critical Invention*; *Project of Perpetual Peace, Comment*; *The Europe of Kant*; *Responsibility of the Modern: On Kantian Universalism*; and several books on political philosophy.

**Chaehyun Chong** is Associate Professor in the Department of Philosophy at Sogang University, South Korea. His areas of expertise are Ancient Chinese Philosophy, Philosophy of Language and Logic, and Confucianism. He recently published "Moism: Despotism or Democrat-

ic?" in *Journal of Chinese Philosophy* (2008) and is working on a project on Mou Zongsan and Kant.

**Predrag Cicovacki** is Professor in the Department of Philosophy, College of the Holy Cross, Worcester, MA.

**David Cummiskey** is Professor of Philosophy at Bates College, Lewiston, Maine. His research focuses on debates between Kantians and consequentialists, and on the cross-cultural study of ethics. He is author of *Kantian Consequentialism* and recent articles appearing in *Morality and Global Justice*, *Metaphilosophy*, *Rethinking Kant—Current Trends in American Kantian Scholarship* and *Utilitas*.

**Adriano Naves de Brito** is Professor in the Department of Philosophy at Unisinos University, Brazil. Besides studying in Brazil, he studied in Germany (Bielefeld and Tübingen). He is a member of the *Kant-Gesellschaft* in Brazil and is President of the Brazilian Society for Analytic Philosophy. His current research is in the field of naturalized ethics.

**Phil Enns** is Lecturer in Philosophy in the Graduate and Islamic Theology Departments, Universitas Islam Negeri Sunan Kalijaga (State Islamic University) in Yogyakarta, Indonesia. He specializes in the relationship between rationality, democracy, and religion, focusing on Kant and Habermas. He has published articles in *The Heythrop Journal* and *The International Journal for Philosophy of Religion*.

**Rainer Enskat** has been Professor of Philosophy at the University of Heidelberg since 1984, and Professor of Philosophy at the University of Halle since 1992. He has written books on *Kant's Theory of Geometrical Objects*, *Truth and Discovery*, *Hegel's Theory of the Practical Conscience*, *Authentic Knowledge*, *Conditions of Enlightenment*, and numerous articles on related themes.

**Wolfgang Ertl** is Associate Professor in the Department of Philosophy/Ethics at Keio University, Tokyo. His research interest is focused on topics at the intersection of metaphysics and ethics and has resulted in books such as *Kants Auflösung der dritten Antinomie* and *David Hume und die Dissertation von 1770*.

**Hans Feger** (PhD, 1995; Habilitation, 2004) is Professor of Philosophy at Free University of Berlin, Germany. His research interests are focused on German Idealism (Kant, Schelling) and on Kierkegaard. His publications include: *Die Macht der Einbildungskraft in der Ästhetik Kants und Schillers* and *Poetische Vernunft. Moral und Ästhetik im Deutschen Idealismus*.

**Mihaela C. Fistioc** (PhD, University of Chicago, 2000) specializes in the aesthetics of Kant and Plato, and the connections between Ger-

man Idealism and classical Greek philosophy. She has taught philosophy at Yale University. She is author of *The Beautiful Shape of the Good* and articles on aesthetic experience and the transcendental unity of apperception.

**Courtney David Fugate** is Professor in the Department of Philosophy at The American University in Cairo, Egypt.

**Klaus-Gerd Giesen** is Professor of Political Science at the University of Auvergne, France, and has been a visiting professor in several universities in France, Belgium, and Switzerland. His numerous publications are in the fields of political philosophy and international relations. See: [www.giesen.fr](http://www.giesen.fr).

**Vasil Gluchman** is Director of the Institute of Philosophy and Ethics and Professor of Philosophy and Ethics at University of Presov, Slovakia. His publications include eight books (most recently: *Ethics in Slovakia—past and present*; *Ethics and Reflection of Morality*; *Human Being and Morality in Ethics of Social Consequences*) and many articles and edited books.

**Ronald M. Green** (PhD, Harvard, 1973) is Eunice and Julian Cohen Professor for the Study of Ethics and Human Values, Dartmouth College, and Director of Dartmouth's Ethics Institute. He has authored seven books, including: *Religious Reason*; *Religion and Moral Reason*; *Kierkegaard and Kant*; and *Babies by Design: The Ethics of Genetic Choice*.

**Robert Greenberg** is Professor of Philosophy at Brandeis University in Waltham, Massachusetts. His research on Kant's theory of knowledge and its relation to work in the philosophy of language has produced two books—*Kant's Theory of A Priori Knowledge* and *Real Existence, Ideal Necessity: Kant's Compromise and the Modalities without the Compromise*—and several papers.

**Robert Gressis** is Assistant Professor at California State University, Northridge. Having completed his dissertation, *Kant's Theory of Evil: An Interpretation and Defense*, at the University of Michigan, he is now working on relating Kant's theory of radical evil to his moral psychology, specifically his conceptions of maxims, happiness, pleasure, and character.

**Natascha Gruber** is Lecturer at the University of Vienna, and Research Scholar in the BBRG Research Group, University of California, Berkeley. After completing her doctoral dissertation, on the notion of the "Transcendental" in Kant and its reception within Analytic Philosophy, her research interests have been in Epistemology, Transcendental Philosophy, Logic, and Philosophy of Language.

**Kiyoshi Himi** (PhD, Kyoto University) is Professor of Philosophy at Suzuka International University, Japan. His articles in English include: “Practical Freedom—‘One from’ or ‘One of the Natural Causes?’”; “Moral Law as the Ratio Cognoscendi of Freedom in Kant’s Philosophy”; and “The Final End of Creation and the Moral Proof of the Existence of God”.

**Anita Ho** (PhD, University of Alberta) is Assistant Professor in the Centre for Applied Ethics at the University of British Columbia. Her articles have appeared in *The Journal of Medical Ethics*, *American Journal of Bioethics*, *Journal of Clinical Ethics*, *Teaching Philosophy*, etc. She is currently writing a book on trust and autonomy in medicine.

**Soraj Hongladarom** is Associate Professor of Philosophy at Chulalongkorn University in Bangkok, Thailand. He has published books and articles on bioethics, computer ethics, and the roles that science and technology play in the culture of developing countries. He is also working on a project on a dialog between Buddhism and science.

**Mohammad Raayat Jahromi** is a PhD student in Philosophy at Allameh Tabataba’i University, Tehran, Iran. He is also a Faculty member at the Institute for Islamic Culture and Thought, Tehran, and the Iranian Association for Philosophy of Religion. His articles include “An introduction to Culturalized Philosophy, playfulness of understanding in Wittgenstein and Gadamer”.

**Bernhard Jakl** is Assistant Professor at the Chair of Private Law, Philosophy of Law and Medical Law, Faculty of Law, at the University of Muenster, and Associate Lecturer in the Philosophy Faculty at LMU Munich. He specialises in interdisciplinary research between jurisprudence and philosophy. His publications include *Recht aus Freiheit (Right from Freedom)*.

**Mohsen Javadi** (PhD, Tarbiat Modares University, Tehran, 1999) is Associate Professor in the Theology and Philosophy Department at the University of Qom, Iran. He specializes in Ethics and in Islamic theology and philosophy. His publications, mostly on Ethics, include five books and multiple articles in the following journals: *Topoi*, *Andisheh Dini*, and *Nameh Mofid*.

**Katsutoshi Kawamura** is Professor at Kwansei Gakuin University, Nishinomiya, Japan.

**Soo Bae Kim** (PhD, Trier University) is Professor of Philosophy at Chungnam National University, South Korea. He has been Visiting Professor in Oxford and Trier. He is a member of the Editorial Board and the Director of the Academic Exchange for the Korean *Kant-Gesell-*

*schaft*. His research focuses on Enlightenment philosophy, ethics, and philosophy of history.

**Takayuki Kisaka** is Professor of Philosophy at Kokushikan University in Tokyo, Japan. His articles in German include: “Innerhalb und Außerhalb der bloßen Vernunft” in: *Wandel zwischen den Welten* (2002) and “Die Möglichkeit der metaphysisch bestimmenden Urteilskraft” (2009). He gave a lecture on Kant’s theoretical philosophy at LMU (München) in January 2010.

**Chong-fuk Lau** (PhD, Heidelberg, 2002) is Associate Professor of Philosophy at The Chinese University of Hong Kong. He has published *Hegels Urteilkritik* (2004) and articles in journals including *Kant-Studien*, *Kant Yearbook*, *The Owl of Minerva*, and *Perspektiven der Philosophie*. He is also a contributor to the new Kant-Lexikon.

**Christine Lopes** (PhD, Birkbeck, London), University of Southampton (2007–2010), founded and runs the Late German Philosophy project ([www.lategermanphilosophy.com](http://www.lategermanphilosophy.com)). Her research interests are in Kant, Early Modern and Early Twentieth Century Philosophies, and Nineteenth Century German Philosophy. She is currently writing a book on Kant’s influence in contemporary moral philosophy.

**Brandon Love** is a Master student in Philosophy of Religion at Trinity International University, Deerfield, Illinois. For most of 2009 he was also an Exchange Research Student in the Religion and Philosophy Department at Hong Kong Baptist University. His thesis focuses on recent interpretations of Kant’s philosophy of religion and their relationship to contemporary theology.

**Ruchira Majumdar** obtained her BA with Honours in Philosophy from Calcutta University and Masters in Philosophy (first in First Class) from Jadavpur University, India. She received multiple awards and research fellowships. After obtaining her PhD from Jadavpur University, on Logic and Philosophy of Language, she has published a book, presented papers, and delivered many lectures.

**Martin Moors** is Professor of Philosophy at the Catholic University of Leuven, Belgium. He has written articles on Kant, Schelling, and Modern Rationality and co-edited *The Concept of Love in 17th and 18th Century Philosophy* (2007).

**Eric Sean Nelson** is Assistant Professor of Philosophy at the University of Massachusetts-Lowell. His publications include articles on ethics, environmental ethics, and the philosophy of nature in Chinese and European philosophy. He is the coeditor of *Addressing Levinas* and *Rethinking Facticity*.

**A.T. Nuyen** is Associate Professor of Philosophy at the National University of Singapore. His articles on Hume and Kant have appeared in journals such as *Hume Studies*, *Kant-Studien* and *History of Philosophy Quarterly*, and on comparative philosophy in *Journal of Chinese Philosophy*, *Asian Philosophy*, *Philosophy East and West*, *Copenhagen Journal of Asian Studies*, and *Dao*.

**Emer O'Hagan** (PhD, University of Toronto, 2001) is Associate Professor of Philosophy at the University of Saskatchewan, Saskatoon, Canada. She works in ethics and meta-ethics. Her recent publications include: "Velleman on the Constitutive Aims of Practical and Theoretical Reasoning" (*Sorites*) and "Animals, Agency, and Obligation in Kantian Ethics" (*Social Theory and Practice*).

**Gregg Osborne** (PhD, Chicago) is Associate Professor of Philosophy at Washington and Jefferson College, near Pittsburgh, Pennsylvania. He previously taught at the American University of Beirut. He has published on both Kant and Hume in journals such as *Kant-Studien*, *Hume Studies*, *The Journal of Philosophical Research*, and *Areté: revista de filosofía*.

**Stephen R. Palmquist** (DPhil, Oxford, 1987) is Professor of Religion and Philosophy at Hong Kong Baptist University. His 100+ publications include nine books (e.g., *Kant's System of Perspectives* and *Kant's Critical Religion*) and multiple articles in *Kant-Studien*, *Faith & Philosophy*, *Journal of Religion*, *Philosophy & Theology*, *Review of Metaphysics*, and *The Heythrop Journal*. See [www.hkbu.edu.hk/~ppp/Kant.html](http://www.hkbu.edu.hk/~ppp/Kant.html).

**Chan Goo Park** (PhD, Tübingen, 1995) is Professor in the Department of Ethics Education, Seoul National University, South Korea. He taught moral philosophy in the Philosophy Department of Hankuk University of Foreign Studies, South Korea, from 1996 to 2003. He has published several books (including textbooks) and multiple journal articles in Korean.

**Peter K.J. Park** is Assistant Professor in the School of Arts & Humanities at the University of Texas, Dallas.

**Wen-berng Pong** (彭文本) is Associate Professor in the Department of Philosophy, National Taiwan University Taipei, Taiwan.

**Alain-Marc Rieu** is Professor of Contemporary Philosophy and Epistemology at the University of Lyon and Research Fellow at the Institute of East-Asian Studies (CNRS), ENS-lsh. His work on the cultural role of knowledge since the 1970s compares philosophical constructions in East-Asia and in Europe. He is author of various articles, studies, and books. See <http://w7.ens-lsh.fr/amrieu/>.

**Nils Röllner**, Professor in the Department of Art and Media at Zurich University of the Arts, holds an MA in philosophy (Freie Universität, Berlin) and a PhD in media theory (Bauhaus University, Weimar). He was Visiting Fellow at ZKM Karlsruhe in 2002–2003. In 2010 he published *Magnetismus—Eine Geschichte der Orientierung*. See [www.romanform.ch](http://www.romanform.ch).

**Marc Rölli** (Dr. phil. habil.) is acting chair of the Philosophy Department at the Technical University, Darmstadt. He is author of *Gilles Deleuze. Philosophie des transzendentalen Empirismus* (English translation forthcoming) and *Kritik der anthropologischen Vernunft*. He is currently co-editing a volume on philosophy and the so-called “spatial turn”. See [roelli@phil.tu-darmstadt.de](mailto:roelli@phil.tu-darmstadt.de).

**Peter Schröder** is Senior Lecturer in Early Modern History at University College London. Main research interest: the History of Political Thought. He has been visiting professor at universities in Paris, Rome, and Seoul. His publications include: *Naturrecht und absolutistisches Staatsrecht. Eine vergleichende Studie zu Thomas Hobbes und Christian Thomasius* (2001) and *Niccolò Machiavelli* (2004).

**Ulrich Seeberg** is Visiting Professor of Aesthetics and Theory of Art at Berlin University of the Arts. He also teaches philosophy at Martin-Luther University, Halle-Wittenberg. His areas of specialization are Kant, German Idealism, Aesthetics, and Phenomenology. His publications include *Ursprung, Umfang und Grenzen der Erkenntnis. Eine Untersuchung zu Kants transzendentaler Deduktion der Kategorien*.

**Susan Shell**, Professor and Chair of the Department of Political Science at Boston College, is author of: *The Rights of Reason: A Study of Kant's Philosophy and Politics*; *The Embodiment of Reason: Kant on Spirit, Generation and Community*; and *Kant and the Limits of Autonomy*. She is co-editing a work on Kant's *Observations and Remarks*.

**Scott R. Stroud** (PhD, Temple University) is Assistant Professor of Communication Studies at the University of Texas at Austin. He works in the areas of pragmatism, comparative philosophy, and Kant studies. His book, *Pragmatism and the Artful Life*, is forthcoming from Pennsylvania State University Press.

**Makoto Suzuki** is a part-time lecturer at Nanzan University and Nagoya University, Japan. He received his BA and MA in letters (moral philosophy) from Kyoto University, Japan, and his PhD in philosophy from The Ohio State University, USA. His research interests include moral philosophy, philosophy of psychology, and history of modern western philosophy.



**Rein Vos** (PhD, University of Groningen, 1989) is University Professor, Department of Health, Ethics & Society, at Maastricht University, The Netherlands. He has served as Visiting Professor at the Centre for Medical Ethics, University of Oslo, and Research Fellow at the Centre for Philosophy of Science, University of Pittsburgh, and at George Washington University.

**Michael Thompson** received his PhD from the University of South Florida. He teaches for the Honors college at USF and the University of Tampa. His research focuses on the role of imagination in Kant's Critical philosophy.

**Bart Vandenabeele** is Professor of Aesthetics and Philosophy of Art at Ghent University. He has published chiefly in philosophical aesthetics, the history of eighteenth- and nineteenth-century philosophy, and the philosophy of language and communication. He is editor of *A Companion to Schopenhauer* and associate editor of *Empedocles: European Journal for the Philosophy of Communication*.

**Stijn Van Impe** obtained his Master Degree in philosophy (summa cum laude, 2007) at Ghent University, where he is currently employed as Fellow of the Research Foundation Flanders. He is preparing a PhD on Kant's different (secular and religious) accounts of a moral community/world. He has presented papers on Kant at several international conferences.

**Mario Wenning** is Assistant Professor at the University of Macau and Humboldt Fellow at the Goethe University of Frankfurt. His research is in the areas of German Idealism, Critical Theory, and Comparative Philosophy. His publications include articles on Adorno, Heidegger, Danto, and Rorty.

**Christian Helmut Wenzel** received a PhD in 1990 in mathematics (algebraic geometry) from the University of Illinois at Urbana-Champaign, and in 1999 a PhD in philosophy, on Kant's aesthetics, from the University of Wuppertal, Germany. He works on Kant, Wittgenstein, philosophy of language, and aesthetics and is now Distinguished Professor at National Taiwan University.

**Ulrich Fritz Wodarzik** is a physicist and philosopher, working as Professor at the University of Applied Science in Worms, Germany. His areas of expertise are German Idealism and Neoplatonism. Selected papers: *Objektivität, Leben und Normativität* (Moskauer Kant-Kongress 2004); *Über die metaphysische Trinität Welt, Mensch und Gott* (X. Kant-Kongress 2005).

**Kwok-Kui Wong** is Assistant Professor in the Religion and Philosophy Department and Humanities Programme, Hong Kong Baptist University, since 2007. He received his doctorate from the University of Tübingen, Germany, in 2001. His interests include Kant, Schelling, Nietzsche, Phenomenology, and Hermeneutics. He has published articles in *Philosophical Inquiry*, *Hegel Jahrbuch*, etc.

**Julian Wuerth** (PhD, University of Pennsylvania, 2000) is Associate Professor of Philosophy at Vanderbilt University. He has written *Kant on Mind, Action, and Ethics* (2011) and co-edited *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics* (2011). He is currently writing *What Should I Do?: Kant's Ethics* and editing *The Cambridge Kant Lexicon*.

**Simon Shengjian Xie** has written a PhD thesis at University of Melbourne, Australia, entitled “Kant as a moral perfectionist—A Critique of contemporary interpretations of Kant’s moral theory and an assessment of it” and published an article in *Kant-Studien*. His main interests are in Kant, rationality, and comparative study between Kant, Laozi, and Confucius.

**Samuel Xia-ling Xie** is Professor in the Sociology Department at Fudan University, since being transferred from the Philosophy Department in 1995, where he taught Chinese Philosophy after obtaining his PhD with *Kant's Sublation of Ontology* (1985). His publications include four books and various articles, including “Aesthetic Judgment: Mental Power to Interpret Confucianism” in *Fudan Journal*.

**Ellen Zhang** holds a PhD in Philosophy of Religion from Rice University (USA). She taught at Temple University before joining the Department of Religion and Philosophy at Hong Kong Baptist University. Her research interests and publications are in the area of Comparative Philosophy and Religion, Daoism (Lao-Zhuang), and Buddhist philosophy (Madhyamika and Chan).

**Günter Zöller** is Professor of Philosophy at the University of Munich. A past Vice President of the North American Kant Society, he also is a member of the Kant Commission of the Berlin-Brandenburg Academy of Sciences. Together with Robert Loudon he has edited Immanuel Kant, *Anthropology, History, and Education* (2007; paperback 2010).



## Index

- Abbott, T.K. 494n, 496n, 718n, 719n  
Abraham 473  
Achenwall, G. 395-6  
Adam 781  
Adams, R. 676, 736  
Adickes, E. 300-1, 475n  
Adorno, T.W. 54-5, 62, 339  
Advaita-Vedanta 12, 194, 203-4  
Alaph 753  
Allais, L. 225-6  
Allison, H.E. 4, 9n, 15, 26, 36-7, 43-8, 101, 110n, 133n, 134, 309n, 311-2, 313-5n, 316-9, 576-7, 641n, 770-1  
Almeida, G.A. de 771n  
Althusius, J. 769  
Alzheimer's 358  
Ameriks, K. 51, 141n, 227, 576n, 679n  
Ames, R.T. 325n, 330n, 652n, 655n  
amphiboly 10, 75, 140-2, 144, 146n, 150, 153-4, 165  
An, Y. 629n  
Angle, S. 591n  
antinomy 11, 23, 74-5, 135, 181-3, 185-6, 188-90, 367, 369, 489, 498, 553-5, 557-9  
apperception 8, 12, 15, 26, 30, 39, 42-3, 45, 48, 56, 103-5, 124, 128, 160-1, 163, 165, 172-4, 179, 188, 194-9, 277, 289-91, 297, 485, 537, 575-6, 578-81, 584, 593-6, 601, 679-80, 688, 709-714  
architectonic 4, 5n, 7, 15n, 25, 564n, 565, 657, 780, 811-21  
Arendt, H. 64, 315  
Aristotle / Aristotelian philosophy 7, 11, 19, 29, 32, 142-3, 148, 169-70, 175, 219n, 254n, 326-7, 332, 401-2, 408, 483, 493-4, 586, 654, 666-7, 696, 733, 737, 741, 801, 812, 814n, 817-8  
Arjuna 715, 719-20  
Aron, R. 446  
Arp, R. 429n  
Asbach, O. 394n  
Asia / Asian philosophy 3-4, 6, 8, 16, 25, 28, 30-5, 53, 66, 69, 288, 335-7, 344, 380, 627-9, 637-8, 654-7, 664, 693, 730, 739, 741-2, 748, 750-53, 755, 758, 761-2, 777, 781-2, 784-5, 780, 820-1  
Atkins, K. 185n, 188n  
Audi, R. 247-9, 255  
Augustine, St. 9, 57, 118-20, 122, 127, 191, 536  
Aune, B. 430n  
Aurobindo, S. 715n  
Austin, J.L. 106n  
autonomy 13, 19, 21, 28, 30-1, 34, 44, 48, 50-2, 60, 74-5, 77-8, 83, 169, 173, 179, 231, 235-6, 238-40, 248, 252, 265-71, 284, 319, 326, 330n, 345n, 349, 351-2, 354, 364n, 367, 371-3, 375-6, 380-1, 383, 386, 391, 401, 405-6, 410-1, 415n, 417-20, 422, 424n, 431, 433, 436-7, 451, 465, 467, 469-75, 477, 484-5, 487, 511-2, 518, 553, 555, 556, 596, 598-9, 617, 620, 635, 649-50, 662, 664-6, 670-3, 678 682, 689, 707, 732-8, 746, 770-2, 775, 791-2, 794, 796-8  
Axinn, S. 353-4n, 365n  
Ayer, A.J. 117n  
Bache, K. 424n  
Bacon, F. 302  
Baczko, L. von 754n

- Badwi, A.R. 735  
 Baron, M.W. 256-7, 732n  
 Baudelaire, C. 747n  
 Bauer, J. 72n  
 Baumanns, P. 367-8  
 Baumgarten, A.G. 158, 161, 221n,  
 273, 274n, 279n, 448  
 Baur, M. 429  
 Bayle, P. 765, 779n  
 Beck, L.W. 439n, 487, 619n, 633n  
 Beckoff, M. 258  
 Beiser, F.C. 566n  
 Beitz, C. 765  
 Bencivenga, E. 218  
 Benedict, Pope 57-9  
 Benjamin, W. 747n  
 Berkeley, Bishop G. 332n  
 Berkhout, S. 370n  
 Bernal, M. 785  
 Bernard, J.H. 443  
 Bernasconi, R. 336n, 344, 788  
*Bhagavad-Gita* 715-23  
 Bickmann, C. 11-2, 194-204  
 Bilgrami, A. 4, 36-7, 46,  
 Black, M. 114n  
 Bobko, A. 22-3, 493-500  
 Böhme, H. 394  
 Bohr, N. 171  
 Boisclair, A. 27, 603-14  
 Boockmann, H. 783n  
 Bormann, F.J. 386n  
 Boxer, C.R. 760n  
 Boyle, M. 696n  
 Bandom, R.B. 107  
 Bretall, R.W. 568n  
 Bretzke, J.T. 632  
 Brook, A. 10, 140-54  
 Brown, D.M. 570n  
 Brugger, W. 384n  
 Bryuschinkin, W. 419n  
 Buber, M. 179  
 Buddha 27, 30, 298n, 604n, 605-6,  
 611-2, 667, 672-5, 685, 687, 695,  
 704n, 705, 707, 712  
 Buddhism; *see Buddhist philosophy*  
 Buddhist philosophy 12, 24, 27, 29-  
 30, 67, 194, 204, 288, 298-9,  
 336-7, 544, 604, 611, 613-4, 654,  
 664-77, 678-91, 695-709, 712  
 Buhle, J.G.G. 779  
 Bunnin, N. 623n  
 Burg, P. 756n  
 Burgelin, P. 404n  
 Burt, E.A. 746n  
 Busche, H. 401n  
 Butts, R.E. 141n  
  
 Caesar 378  
 Champion, C.T. 562n  
 Carnap, R. 653n  
 Cartesian; *see Descartes*  
 Cassirer, E. 300  
 Castillo, M. 20, 438-46  
 categorical imperative 7, 13-4, 17-20,  
 22, 36, 57, 59-61, 63-5, 67-8, 74,  
 78, 93, 198, 234-40, 242, 245-9,  
 263n, 281-2, 287, 291-2, 330,  
 349, 364, 371, 378, 384-7, 394,  
 407, 409, 411, 416, 419, 425-6,  
 429, 431, 433-4, 477, 488-90,  
 555-6, 610, 619, 662, 671-2,  
 716-7, 721, 725, 729, 731, 736,  
 762, 771, 794  
 categories 7-8, 10, 37, 40, 44, 51,  
 100-5, 124, 140-1, 146-8, 154,  
 158, 205, 266, 289, 336, 359,  
 450, 470, 491, 526, 528, 545-6,  
 583-4, 593, 595, 597n, 616, 658,  
 660, 685, 706, 812, 814-21  
 Catherine the Great 467  
 Catholic 335, 337, 429n, 525, 530-3,  
 760-1  
 Cato 377-8  
 causation 9-10, 12-5, 18, 23, 44, 102-  
 3, 105, 111-4, 116-7, 127, 160-1,  
 173-4, 198-9, 203, 207, 209, 211,  
 217-30, 235, 244, 246, 265-72,  
 274-5, 278-82, 290-1, 294-6,  
 314n, 340-1, 344, 355, 359-60,  
 366-7, 387-8, 403, 414, 417n,  
 436, 444-5, 449-50, 470, 473,  
 487, 498, 511-2, 525, 528, 532,  
 552-5, 558, 586, 589, 598, 600,  
 620, 642, 645, 658, 667, 670,

- 682, 684, 689-90, 703, 705, 707, 736, 772, 789, 791, 810
- Caygill, H. 268n
- Ch'eng, H. 634
- Chan, W.C. 764n
- Chan, W.T. 630-8n
- Chang, T. 636-7
- character 12, 20-1, 32, 47, 67-9, 74, 82, 86n, 99, 107, 194, 203, 234-7, 239, 242, 245-6, 247n, 250n, 251, 252n, 263, 266-7, 269, 274-8, 280n, 282-3, 293, 315n, 327, 345n, 350, 356, 383-5, 389, 403, 406-8, 416, 435, 447-54, 468, 479, 494, 521, 545, 555, 569-70, 631-4, 637, 650, 663, 666, 668, 670, 676, 696-7, 733, 749, 767, 770n, 782, 789, 811
- Chauvier, S. 762
- Chen, S.B. 627
- Cheng, A. 748n
- Cheng Brothers 86
- Cheng, C.Y. 4, 6, 7, 27, 74-96, 332n, 591n, 615-6, 623-4, 633-4
- Chinese philosophy ..., 4, 7, 25-6, 33, 35, 335, 337, 581n, 595, 603, 614, 623n, 630n, 653n, 656, 748, 764-6, 768-72, 775-6, 800-1, 806-7
- Ching, J. 764n, 765n, 767-8, 769n, 770, 775n
- Chisholm, R.M. 111n
- Chiu, H.H. 761n
- Chong, C. 27, 615-24
- Chou T.I. 632, 634
- Christian / Christianity 22, 25, 31, 55-9, 90, 280, 299, 337, 362, 472, 485, 494, 529-33, 535, 562-3, 569-72, 603, 624, 654, 724-7, 730, 745-6, 749n, 751-2, 777
- Chu Hsi 631-2
- Chuang Tzu; *see* *Zhuangzi*
- Chun, H.L. 591n
- Cicovacki, P. 22, 485-92
- civilization 214, 419, 438, 440-1, 451, 453, 757-8, 760, 761n, 782, 784-5, 787-90
- Clarke, J.J. 657
- Cleary, T. 71n, 686n
- Colbert, J.B. 759
- Collins, S. 667n
- community 3, 13-5, 17, 19-20, 22-3, 33, 76, 94, 291, 308, 310n, 314-5, 320, 322, 354n, 356, 364n, 365, 367, 380, 399, 409, 425n, 428, 430, 432n, 433, 436-7, 441, 457, 463, 498, 500, 529, 534, 558, 561-2, 567, 623, 639, 648, 651, 652n, 662, 667-8, 690, 752, 756, 762-3
- Confucianism; *see* *Confucian philosophy*
- Confucian philosophy 6-7, 11, 16, 18, 27-8, 32, 35, 74-96, 182, 328-9, 337-8, 370, 375, 380-1, 591, 603-4, 613-4, 615-24, 627-38, 645, 656-7n, 667-8, 675, 741-52, 765-6, 768-70, 775, 806
- Confucius 6, 16, 28, 33, 53, 66-9, 76-7, 81-5, 87, 90-3, 95, 182, 186, 245n, 321-32, 373, 377, 419n, 629-30, 632, 645, 655n, 764-76
- Cook, T. 757
- Cooley, D. 372n, 376n, 377n
- Copernican Perspective 25, 564-5
- Copernican revolution 56, 107, 524, 543, 565, 616, 656, 658, 742-3, 747, 802
- Copernicus, N. 56
- Craig, E. 241-2, 674n
- Crawford, D. 309, 317, 319n
- Crusius, C.A. 14, 273, 277, 280-3
- Cullberg, J. 179
- cultivation 5, 8, 13, 16-7, 19-20, 25, 28, 32, 34-5, 79-82, 87, 91-2, 97, 193, 322n, 326n, 338, 340, 342, 346-7, 356, 360, 364n, 375, 440, 452, 471, 509, 573, 625, 631-2, 634-5, 637, 639-40, 644-50, 652, 663, 703n, 719, 748-50, 770n, 782, 785, 821
- Cummiskey, D. 29, 373, 664-77
- Custance, G. 300n
- Dalai Lama 668n, 669, 673
- Dao / Tao 6, 11-2, 16-7, 28-9, 35, 62, 69, 71-3, 83, 86, 92, 95, 182,

- 187, 191-2n, 194, 204, 321, 329-34, 336-9, 342n, 344-6, 624, 630-3, 635-6, 637n, 645, 648, 653-6, 657n, 659-63, 806, 813-4
- Daoist philosophy 6, 11-2, 16-7, 28-9, 35, 69, 71-2, 182, 192n, 194, 204, 333-4, 336-9, 342n, 344-6, 653-63, 807, 814
- Darraz, A.A. 735
- Darwall, S. 667
- Davidson, D. 217, 224n
- De Brito, A.N. 14, 265-72
- De Robespierre, M. 62n, 63, 742
- Decher, F. 520n
- deism 533, 742
- Delhom, P. 394n
- Deng, X. 54
- Derathé, R. 404n
- Dermigny, L. 760n
- Descartes, R. 6, 46, 55, 90, 181, 185-6, 188, 212n, 403, 485, 524, 594, 679, 684, 695
- Devil; *see Satan*
- Devolder, K. 361n
- Dilthey, W. 783n
- Diogenes 180
- Diogenes Laertius 785
- Dorff, Rabbi E. 369
- Dougherty, F.W.P. 788
- Dreyfus, H.L. 59n
- Dreyfus, S.E. 59n
- Droysen, J.G. 783
- Dryer, D.P. 132n, 133n
- Du, S. 685-6
- Dubs, H.H. 646n
- Eberhard, J.A. 781
- Eichmann, A. 64
- Elberfeld, R. 661n
- Ellington, J.W. 256n, 487n
- Elliott, R.K. 317
- embryo; *see fetus*
- Emerson, R.W. 734
- emotion 29, 63, 65, 69, 72, 93, 113, 249n, 258, 351, 356, 373, 375, 490, 649, 667-71, 673, 677, 696, 703n, 788, 805
- emptiness 11, 24, 29-30, 60, 65, 71-2, 80, 95, 105, 121, 132n, 149, 188-9, 192, 200, 279, 281, 542-4, 546, 578, 580, 654, 656, 660, 662, 667, 669, 682, 684-7, 699, 724
- Engel, S.M. 536n, 537n, 539n, 540n
- Enlightenment 16, 33, 55-6, 58-9, 62, 282n, 304, 326, 335-6n, 338-9, 383, 398, 439-40, 459, 523-4, 533, 544, 549, 606, 612, 635, 637, 646, 675-6, 678, 683, 690, 712, 731, 765n, 769, 772-4, 781, 783, 787-8
- Enns, P. 21, 455-62
- Enskat, R. 13-4, 233-46
- Epicurean philosophy 27, 604, 606-9, 614
- Ertl, W. 12-3, 217-30
- Esser, A. 385n, 386n
- euthanasia 18, 370-1, 378-80, 804
- Euler, J.A. 301
- evil 5, 14, 23, 28, 58, 64, 214n, 242, 268, 276, 303, 314, 371, 440, 451, 453, 478, 493, 499-500, 501-8, 512-8, 521-2, 527, 529, 534, 546, 559-62, 566, 569, 571-2, 606, 613, 632, 636-7, 639-48, 673, 698, 699, 715, 758
- Ewing, A.C. 129n
- Falkenstein, L. 149n, 150n, 227
- Fan, R. 375n, 381n
- Fazhang 604n
- Feger, H. 7, 23, 511-22
- Feldman, S. 353, 364n
- Feng, Y. 72
- fetus 17, 349-5, 357, 363-9
- Feuerbach, L. 179
- Fichte, J.G. 169, 170n, 176, 177n, 178n, 235n, 283, 388n, 391-2n, 513, 600
- Fingarette, H. 322n, 330n
- Firestone, C. 563, 569n, 571
- Fischer 753
- Fischer, N. 228n
- Fistioc, M.C. 26, 332n, 585-91
- Formisano, M. 394n
- Foroughi, M.A. 734

- Foucault, M. 61, 744n, 748n  
 Frederick the Great 769  
 Frege, G. 114-5  
 Freud, S. 275, 628  
 Fricke, C. 296n  
 Friedman, M. 227, 301  
 Friedrich William II 471n  
 Fromm, S. 536n  
 Fugate, C.D. 14, 273-84  
 Fuhrmans, H. 521n  
 Funke, G. 319n  
 Furlus, E. 300n
- Gandhi, M. 250  
 Gardner, S. 214  
 Garfield, J.L. 683n  
 Geach, P. 114n  
 Geier, M. 536  
 Geismann, G. 354, 392n  
 Genova, A.C. 319n  
 Gentzler, J. 374n, 379n  
 George III, King 754  
 George, R.P. 363n  
 Giesen, K.G. 33, 753-63  
 Gilbert, W. 302  
 Ginsborg, H. 296-7n, 310-2  
 Glasenapp, H. von 765n, 766, 767n  
 Glissant, E. 740  
 globalization 20, 35, 438  
 Gluchman, V. 13-14, 256-64  
 God 6, 12, 22-4, 26-7, 31-2, 55-9, 62-3, 74-5, 90, 164, 167, 169, 172, 174n, 213, 217, 220-1, 230, 253, 277, 282n, 336-7, 346, 402, 424, 466, 469-70, 473, 476, 478-83, 498-500, 505, 516-20, 525-41, 543, 545-9, 551, 557, 559-61, 565-7, 569-72, 585, 587-8, 590-1, 598, 603, 609-14, 617, 624, 636, 658, 662, 699, 715, 718-20, 722-4, 726n, 727-30, 734-8, 742-3, 745-6, 751, 765, 771, 785, 801, 804  
 Godwin, W. 250n  
 Goh, C.T. 627  
 Golden Rule 6, 7, 66-8, 77, 79-80, 95, 245n, 419, 421  
 Goldman, A. 111
- Gomez-Lobo, A. 363n  
 grace 114, 278n, 424, 505, 527-8, 566, 571  
 Graham, A.C. 71n, 333n  
 Gray, J. 62  
 Green, J. 753, 756  
 Green, R.M. 17, 349-57, 364n  
 Greenberg, R. 9, 109-17  
 Greene, T.H. 486n  
 Gregor, M.J. 221n, 223, 305n, 486n  
 Gressis, R. 23, 501-8  
 Grice, H.P. 111n  
 Grier, M. 211n  
 Griffin, J. 249n  
 Grotius, H. 769  
 Gruber, N. 17-8, 358-69  
 Guenin, L.M. 350n  
 Guimarães, L. 272n  
 Gunderson, M. 378n  
 Günther, G. 169, 170n, 171n, 175-7n, 179  
 Guo, Z. 627  
 Guyer, P. 44n, 109-10, 129n, 141n, 305n, 309-12, 426-7, 429-31, 433n, 434, 566n, 590n, 650n
- Habermas, J. 21, 53, 59, 73n, 388n, 455-8, 461  
 Hall, D.L. 655n  
 Halley, E. 301  
 Hamacher, W. 472  
 Hamann, J.G. 54  
*Hamlet* 630  
 Han, J.K. 543n  
 handedness 31, 150-1, 727-30  
 Hanna, R. 206, 226  
 Hansen, C. 6, 322n, 654n, 661  
 happiness 20, 27, 203, 250n, 253, 291-2, 299, 374, 430, 443-5, 473, 493, 495-6, 498, 501, 505, 507, 533, 541, 548, 556-7, 591, 600-1, 603-4, 606-12, 614, 639, 642, 644-5, 650, 660, 675, 716-9, 722, 728-30, 733, 769-70  
 Hare, R.M. 429, 431, 676, 734  
 Hartman, R.S. 537n  
 Hartmann, N. 22, 485, 489-92  
 Harvey, P. 676n



- Hay, G. 754  
 Heaven 7, 61, 75, 77, 82-4, 87, 89-91, 303-4, 328, 334, 345-6, 467, 472, 596, 624, 630, 636-8, 673, 765, 820  
 Hegel, G.W.F. 32, 54, 60-2, 125, 169, 180n, 235n, 239n, 393n, 628, 656, 747-52, 768, 775  
 Heidegger, M. 123-5, 126n, 190-1, 444, 520, 520-1n, 653n, 654, 743, 744n  
 Heim, K. 179  
 Heimsoeth, H. 174n  
 Heine, H. 742-3  
 Heintel, P. 536n  
 Heisig, J.W. 654n  
 Henningfeld, J. 520n  
 Henrich, D. 36, 49, 289n, 467, 475  
 Herdegen, M. 384n  
 Herder, J.G. 766n, 775, 783n, 784n, 796n  
 Herman, B. 424n  
 Herrmann, C. 521n  
 Hershock, P.D. 652n  
 Hesychius, 753  
 Heubel, F. 748  
 Hill, T.E. 424n, 666  
 Himi, K. 24, 550-62  
 Himmelmann, B. 415n  
 Hindu philosophy 12, 31, 67, 194, 203-4, 719  
 Hinduism; *see Hindu philosophy*  
 Hintikka, J. 106  
 Hirsch, A. 394n  
 Hitler, A. 63  
 Ho, A. 18, 370-81  
 Hobbes, T. 340, 391-3, 402, 455, 457, 534  
 Hoche, H.U. 419n  
 Höchsmann, H. 333n, 344n, 345n  
 Höffe, O. 297n, 429, 763n  
 Holtman, S. 424n, 426n, 436  
 Hongladarom, S. 30-1, 709-14  
 Honneth, A. 239n, 446  
 Horatius 753  
 Horkheimer, M. 339  
 Hsai, A. 770n  
 Huainanzi 53, 66, 69-71  
 Huayan Buddhism 29, 604, 678-91  
 Hudson, H. 217-8  
 Hudson, H.H. 486n  
 Hui, E. 375n, 377n  
 humanity 13, 18, 21-2, 34-5, 63, 67-8, 74-6, 81-3, 85-88, 91-3, 95-6, 190, 248, 254, 256-63, 302, 308, 331, 349-50, 354-5, 364, 371, 373, 377, 380, 384, 387-8, 413, 426-7, 430-1, 437, 440-2, 445, 447, 452-4, 468, 472-3, 478-80, 485-6, 491, 500-1, 525, 528, 534, 565-72, 610, 623, 633-4, 641-2, 649, 664-6, 707, 719, 726, 730, 736-7, 745, 747, 750, 752, 769, 806  
 Hume, D. 12, 14, 65-6, 181, 185, 187-8, 193, 204, 266, 272n, 303, 314n, 316, 523-5, 536, 667, 742, 745  
 Hünning, D. 391n, 394n  
 Hunter, I. 393n  
 Hurley, P. 666  
  
*I Ching (Book of Changes)* 7-8, 811-21  
 Ideal 9, 12, 13, 16, 20, 29-31, 35, 62, 89, 110-3, 116-7, 119, 125, 135, 138, 153, 157, 169-70, 175, 199, 201, 281, 220, 223, 228, 247, 249, 251, 255, 264, 270, 298, 307, 332, 402, 481, 482, 512-3, 519n, 520, 553, 558-9, 561, 622-3, 655, 657, 659, 681, 712, 727-9, 799, 802, 806  
 Ilkhani, M. 531n, 532n  
 imagination 9, 12, 35, 78, 95, 122-5, 127, 141, 160, 181, 187, 196, 205, 208-14, 216, 300, 302-3, 305, 311, 316-7, 319, 326, 328-9, 338, 341, 345, 445, 476-7, 514, 578, 712, 801, 805-6  
 immortality 10, 31, 74, 159, 164, 166, 212n, 213, 346, 367, 402n, 465, 469-70, 546, 551, 557, 559-60, 565, 572, 598n, 658, 718, 720, 722, 724, 726, 729-30, 801  
 inclinations 6, 23, 28, 63-5, 84, 268, 291-2, 301-2, 311-2, 315, 372n,

- 374, 397, 407-8, 410-2, 415-6,  
449-54, 466-7, 469, 486, 495-6,  
498-9, 501-4, 506-8, 592, 619,  
630, 636, 639-45, 647-9, 662,  
698, 703, 707, 771, 795, 797
- Infield, L. 408n, 476n
- intellectual intuition 25-7, 35, 220,  
559, 573, 580-5, 591, 593, 595-6,  
610-1, 614, 616-24, 743, 748,  
805-6
- intersubjectivity 72, 320, 339
- Isfahani, M.H. 737n
- Islam 57, 67, 673n, 732-8
- Isokrates 66
- Ivanhoe, P.J. 7, 322n, 330n, 332n,  
646n
- Jacobi, F.H. 468-9, 484, 519
- Jacobs, N. 571n
- Jahromi, M.R. 23-4, 523-35
- Jakl, B. 18, 382-90
- Javadi, M. 31-2, 732-8
- Jesus 6, 58, 478, 527, 561-2, 570-1
- Jews / Jewish 57, 369, 529-30, 718,  
779, 781, 786
- Job 526
- Johnson, R.N. 432n
- Jones, H.S. 589n
- joy 86, 113, 280, 607, 609, 674
- Joy, C.R. 550n
- Kagan, S. 671n
- Kamm, F. 666
- Kant, I. *passim*
- Kantorowicz, E. 306n
- Karam, Y. 735-7
- karma 669-70, 675-6, 689, 715, 720-  
23
- Kasulis, T. 690
- Katz, S. 681n
- Kawamura, K. 19-20, 415-23
- Kegley, C.W. 568n
- Kelsen, H. 397n
- Kemal, S. 319n
- Kemp Smith, N. 109, 121n, 129n,  
140n, 141n, 143n, 187n, 195n,  
680n, 710n, 800n
- Kersting, W. 397n, 406n
- Kierkegaard, S. 178n, 512, 518-20
- Killen, A.R. 568n
- Kim, S.B. 34, 791-8
- kingdom of ends 20, 218n, 257n, 271,  
356, 424-5n, 426, 428, 432, 434,  
451, 470, 486n, 638-9, 649, 719,  
723
- Kirkland, R. 337n
- Kisaka, T. 31, 724-31
- Kitcher, P. 4, 5, 26, 36-52, 679n
- Kleingeld, P. 447-8, 793n, 794, 795n
- Kline, T.C. 646n
- Kneller, J.E. 353n, 364n
- Kolak, D. 187n
- Kongzi; *see Confucius*
- Körber, H.G. 305n
- Körner, S. 532n
- Korsgaard, C. 181, 187n, 189-90,  
218, 253-4, 256-7, 424n, 437n,  
486, 666
- Krishna 715-6, 719-23
- Kubin, W. 67n
- Kuhn, M. 471n
- Kulenkampff, J. 297n
- Kwan, T.W. 5, 52n
- Lacan, J. 744
- Langton, R. 225-7, 228n
- Lao Tzu; *see Laozi*
- Laozi 11, 182n, 187, 337, 657, 659-  
60, 806, 812
- Larrimore, M. 788, 790
- Las Casas, B. de 758
- Lau, C.F. 8-9, 99-108
- Lavater, J.K. 484n
- Laveaux, J.C. 783
- Le Blanc, C. 70
- Lee, J. 649-50
- Lee, M.H. 748
- Legge, J. 68, 69n, 770n
- Lehmann, G. 418-9n
- Leibniz, G.W. 14, 106n, 140-5, 151-  
3, 171, 181, 194, 219, 223, 225,  
273, 277-81, 283, 332, 335-6,  
340, 403, 654, 745, 765-6, 768-9,  
775, 777, 781
- Lequan, M. 435
- Lévi-Strauss, C. 744n

- Li, A. 636-7  
 Li, C. 637  
 Liang, T. 84n  
 Lichtenberg, G.C. 401n  
 Liddell, H.G. 589n  
 Liezi 71  
 Lillie, W. 718n  
 Lin, Y. 632  
 Linrothe, R. 673n  
 Linthe, H.J. 305n  
 Locke, J. 11-2, 42-3, 102, 140-3,  
     181n, 185-6, 194, 204, 402, 536,  
     679, 745  
 Long, A.A. 607n  
 Longuenesse, B. 129n, 133n, 315n  
 Lopes, C. 12, 205-16  
 Lotter, F. 783n  
 Louden, R.B. 650n, 703  
 love 22, 34, 48, 58, 77, 85, 88-9, 96n,  
     257, 330n, 346, 370, 372-5, 381,  
     446, 477, 486, 490-1, 501, 505,  
     507, 518, 630-2, 641, 644, 647-  
     51, 672, 717, 805,  
 Love, B. 25, 563-72, 819n  
 Lown, G. 820  
 Lucretia 377-8  
 Luetkehaus, L. 653n, 657n  
 Lusthaus, D. 333n, 674n  
 Luther, M. 525  
 Lyotard, J.F. 315n, 339n  
  
 Macdonald, G.F. 117n  
 Machiavelli, N. 391, 394n, 402, 766,  
     774  
 MacIntosh, J.J. 106n  
 MacKay, D.M. 175n  
 Mainusch, H. 62n, 64n  
 Mair, V.H. 71n  
 Majumdar, R. 31, 715-23  
 Makkreel, R. 341n, 342, 343  
 Malebranche, N. 336  
 Manchester, P. 813n, 814  
 Manninen, B.A. 17, 349-50, 353,  
     354n, 359, 361, 366-7, 369  
 Maritain, J. 735  
 Marshall, P.J. 759n  
 Martin, R. 187n  
 Masaryk, T.G. 264  
  
 Mathieu, V. 482, 483n  
 Matthews, E. 590n  
 McCarty, R. 506n  
 McCloskey, M.A. 424n, 427, 429  
 McCrudden, C. 382n, 383n  
 McDowell, J. 206, 215-6  
 McGeer, V. 695n  
 McMahan, J. 362n  
 McRobert, J. 154n  
 Medicus, F. 792n, 797n  
 Meier, G.F. 37-8  
 Meiners, C. 34, 782-90  
 Melnick, A. 429  
 Mencius 27, 53, 63, 66, 68-9, 80, 82,  
     86-7, 622, 638, 646  
 Mendelssohn, M. 402n  
 Mengzi; *see* *Mencius*  
 Merkel, R. 394n, 397n  
 Merleau-Ponty, M. 190-3  
 Michalson, G.E. 486n  
 Modaresi, M.R. 735  
 Moddares, A.A. 734  
 Mohist philosophy 6  
 Mohr, J.C.B. 550  
 Montaigne, M. de 524  
 Montesquieu, C.L. 404  
 Moore, J.M. 425n, 430n  
 Moors, M. 22, 475-84  
 Morgenstern, C. 778  
 Moses 781  
 Mother Teresa 250  
 Motherby, R. 754  
 Mou, B. 662n  
 Mou, Z. 7, 25-8, 573-624, 748, 764n  
 Müller, M. 764n  
 Murdoch, I. 485, 492  
  
 Nagarjuna 29, 682-4  
 Nagasena 683-4  
 Nagel, T. 274, 666, 695  
 Nagl, L. 536n  
 Nehamas, A. 181  
 Neiman, S. 36  
 Nietzsche, F.W. 55, 65, 181, 193,  
     306-7, 337, 479, 654, 750n, 764,  
     812  
 Nelson, E.S. 16-7, 333-46  
 Nelson, H.L. 371n, 372n

- Neville, R. 660  
 Newton, I. 301, 331-2, 523, 539-40  
 Newtonian physics 93, 121, 227n,  
 301, 327, 523-5, 540  
 Nhat Hanh, T. 298n, 672n  
 Nirvana 606, 654  
 nothingness 28, 337, 479, 482, 517,  
 653-7, 659-63  
 Nozick, R. 666, 805n  
 Nussbaum, M. 379n  
 Nuyen A.T. 28, 30, 627-38
- O'Hagan, E. 30, 695-708  
 O'Neill, O. 36, 350n, 352n, 363n  
 Obama, President B. 250, 577, 579  
 Ockham, W. of 123, 531  
 orientation 11-2, 15, 22, 28, 191-2,  
 199, 262, 285, 300-7, 330n,  
 334n, 465, 475, 647-52, 669-70,  
 727-9  
 Osborne, G. 10, 129-39  
 Oshima, H.H. 71n  
 Otto, R. 569n, 690  
 Ovid 335, 337  
 Oxtoby, W.G. 765n, 767-8, 769n,  
 770, 775n
- Palmer, M. 333n, 344n  
 Palmquist, S.R. 3-35, 64n, 147n,  
 166n, 300n, 307n, 337n, 364n,  
 370n, 437n, 563-5, 571, 725,  
 657n, 726n, 764, 811-21  
 Pandora 62, 186  
 Pang-White, A.A. 339n  
 paralogsims 10-1, 102-3, 118, 141n,  
 146, 153, 157-61, 163, 165-6,  
 171, 177, 194, 538, 581, 583,  
 679n  
 Parfit, D. 181, 187n, 667n  
 Park, C.G. 24, 536-49  
 Park, P.K.J. 33-4, 777-90  
 Parkinson, G.H.R. 141n  
 Parmenides 653  
 Parret, H. 310n  
 Pascal, B. 69, 536, 702  
 Patzig, G. 241-2, 244  
 Paul, G. 67  
 Peetz, S. 518n
- Penelhum, T. 106n  
 Pereboom, D. 141n, 146n  
 Perry, C. 371n, 378n  
 personhood 3-35, 55, 57-8, 97, 118,  
 128, 130, 155, 194-5, 199, 201,  
 203, 231, 239, 271, 287-8, 291,  
 297-8, 336, 347, 352-5, 357,  
 358-61, 370, 372-3, 375n, 376-  
 81, 385, 389, 425, 433, 437, 438,  
 465, 472, 493, 500, 509, 561-2,  
 563, 566-7, 573, 678-84, 688,  
 690-1, 715-9, 723, 724-5, 728-  
 31, 751, 765-6, 811-3, 821
- Peterman, J. 332n  
 Peters, M. 397n  
 Pettit, P. 257n, 732n  
 Pfothenauer, H. 783n  
 Pietism 90, 280, 525, 533  
 Pistorius 468  
 Pitcher, G. 536n, 537n  
 Plato 19, 180, 401-2, 404, 408, 413-4,  
 585n, 769, 775  
 Pluhar, W. 3, 109, 140n, 143n, 596n,  
 597n, 819n  
 Pogge, T. 434n, 756  
 Pong, W.B. 26, 575-84  
 Potter, N. 699n  
 prayer 58, 483, 527  
 Price, H.H. 111  
 Promies, W. 401n  
 Protestant 525, 533, 757n, 761  
 Pufendorf, S. von 769  
 Pye, M. 674n
- Qin, X. 333n  
 Qui, R. 371n, 375n, 381n
- Radhakrishnan, S. 715n  
 Railton, P. 676  
 Ramachandran, V.S. 72  
 Rand, A. 96  
 Rasmussen, D.M. 59n  
 rational psychology 10, 102, 158-9,  
 166, 343, 594  
 rational theology 158, 530-3, 535  
 Ratzinger, J.; *see Benedict, Pope*  
 Rawls, J. 21, 397n, 424n, 429, 455-8,  
 461, 665, 734, 756n

- Reagan, T. 252n  
 Reath, A. 424n  
 Rees, G. 302n  
 Reich, K. 402n  
 Reihman, G.M. 764-6n, 767  
 Reinhold, C.L. 273, 283  
 Reinhold, K.L. 777  
 Reiss, H.S. 766, 769n, 771n, 772, 775n  
 religion 3-6, 17, 21-5, 27, 32, 48-50, 57-59, 61, 90, 170, 198, 203, 214n, 268n, 298-9, 304, 335n, 336, 341, 344, 346-7, 351, 362, 419, 424, 441-2, 455-63, 468, 470-2, 475-84, 486, 499-502, 505, 509, 514-5, 519, 522-3, 525-33, 535, 538, 541, 546n, 548, 550-2, 556-72, 640, 651, 657, 681n, 690, 698, 703, 708, 715, 719, 723, 725-8, 731-3, 737-8, 742-4, 746-7, 757, 765-9, 771, 804, 811n, 814, 818-9, 821  
*ren* (benevolence) 6-7, 67, 69, 74-8, 81-93, 95-6, 321-4, 326-7, 331, 596, 629, 634, 650  
 Rezaei, M. 736n  
 Ricoeur, P. 181  
 Ridley, M. 258n  
 Rieu, A.M. 32, 741-52  
 Rink, F.T. 777n  
 Rischmüller, M. 540n  
 Ritter, J. 419n  
 ritual 16, 94n, 321-6, 328-32, 462, 526, 533, 639, 648-9, 651-2, 675, 785  
 Robber Zhi 652  
 Robespierre, M. de 62-3, 742  
 Robinson, R.H. 683n  
 Rodis-Lewis, G. 607n  
 Roetz, H. 67, 330n  
 Rohden, V. 172n, 499n, 652n, 772n  
 Röller, N. 15, 300-7  
 Rölli, M. 20-1, 447-54  
 Rorty, R. 181, 188  
 Rosemont, H. 322n, 325n  
 Roser, A. 302n  
 Ross, D. 13, 247-51, 253, 255  
 Ross, W.D. 235n, 493n  
 Rousseau, J.J. 67-8, 392n, 402, 404, 405n, 453, 756  
 Rueger, A. 329n  
 Ruffing, M. 772n  
 Rupp-Eisenreich, B. 783  
 Ruse, M. 369n  
 Ryan, M.L. 483, 484n  
 Sagan, A. 363n  
 Sagoff, M. 354n  
 Santinello, G. 778n, 783n  
 Santozki, U. 402n  
 Sartre, J.P. 53  
 Satan 478, 516  
 Saunders, D. 393n  
 Scanlon, T. 666  
 Schaper, E. 658n  
 Scheffler, S. 665n  
 Schelling, F.W.J. 23, 125, 235n, 452, 511-22  
 Schiller, F. 263, 332, 734  
 Schings, H.J. 783n  
 Schmitt, A. 401n  
 Scholz, H. 475n  
 Schopenhauer, A. 59, 66, 537, 539  
 Schrei, H.H. 419n  
 Schröder, P. 18-9, 391-8  
 Schröer, C. 386n  
 Schulz, W. 178n  
 Schwan, A. 467n  
 Schwarz, B.I. 322n  
 Schwarz, M. 416n  
 Schwarz, W. 537n  
 Schweitzer, A. 24-5, 550-1, 553, 557-62  
 Scott, R. 589n  
 Seeberg, U. 15, 287-99  
 Seebohm, T.M. 319n  
 self-cognition 8-9, 99, 103, 105, 108  
 self-consciousness 14, 26, 29-30, 42, 92, 103, 118, 212-3n, 276-7, 281, 283-4, 378n, 380-1, 543, 575-7, 580-3, 594-6, 601, 679-82, 688, 709-11, 713  
 self-legislation 19, 32, 77, 238, 436-7, 733-4, 736  
 self-love 48, 370, 477, 486, 501, 505, 507, 641, 644, 649-51

- Sellars, W. 43-44, 50-1, 106, 215-6  
 Sensen, O. 386n  
*sensus communis* 16, 308, 313-5, 317-8, 321, 341-3  
 Seth, J. 716  
 Severus, A. 66  
 Seyss-Inquart, A. 63  
 Shakespeare, W. 630  
 Shell, S. 21-2, 465-74  
 Shen, M. 371n, 273n  
 Sherman, N. 7  
 Shieh, S. 591n  
 Shien, G.M. 653n  
 Shun, K.L. 326n, 332n  
 Sicha, J.F. 44n  
 Siderits, M. 704n  
 Sidgwick, H. 273, 283  
 Siep, L. 239n  
 Sim, L.J. 632  
 sincerity 28, 30, 82, 89, 344, 627-38, 700  
 Singer, P. 363n, 666  
 Sinnott-Armstrong, W. 249n  
 Slote, M. 257n, 732n  
 Smith, A. 757  
 Sobel, J.H. 301n  
 Solomon, R. 690-1  
 Sömmerring, S. 163  
 soul 10-2, 35, 63, 74, 141n, 157-64, 166, 172, 179, 182, 212-3n, 229, 277, 322, 343, 367, 401-2, 450, 452, 478, 485, 520, 543, 546, 551, 556-7, 559, 597n, 617, 659, 718, 720, 722-4, 726, 729-30, 736, 805  
 space 3n, 83, 100-1, 106-7, 111, 120, 124, 126-8, 135, 148, 150-2, 161, 163, 183, 186, 194-5, 198-9, 201-3, 205-7, 212n, 217, 220, 223, 265, 287-90, 294, 296, 299, 324, 330, 334, 345n, 385, 403, 418, 476, 496, 502, 519, 541, 567, 660, 687, 695n, 727, 744  
 Spaeth, H. 741n  
 Spinoza 31, 161, 181n, 184n, 336-7, 340, 344, 346n, 727  
 Spittler, L.T. 789n  
 Stark, W. 408n  
 Starke, F.C. 448n  
 Stegmaier, W. 64n  
 Stoics / Stoicism 27, 401-2, 435, 437, 477, 491, 604, 606-9, 614, 771n, 775  
 Strauss, L. 741  
 Strawson, P.F. 12, 110n, 112, 117n, 129n, 218, 219n, 223  
 Strong, J.S. 667n  
 Stroud, S.R. 28, 639-52  
 Sturma, D. 179n  
 subjectivity 10-1, 14, 16, 20, 22, 139, 167, 170-180, 185n, 273-6, 278-81, 283-4, 310n, 339n, 404, 413-4, 471, 511, 518, 543, 615, 624, 679, 690, 732  
 suffering 30, 66, 69, 259-60, 268, 351, 371-2, 374, 376, 379, 386, 444-5, 503-4, 607-8, 654, 664, 666, 668-9, 671, 673, 675, 697, 703-8  
 Sun, Y. 333n  
 Suzuki, M. 13-4, 247-55  
 Swanson, P.L. 605  
 Swedenborg, E. 335, 337  
 Tao, J. 381n, 683n  
 Taoism; *see Daoist philosophy*  
 Taylor, C. 756  
 Tennemann, W.G. 34, 777, 779-82, 784, 787, 790  
 Terra, R. 652n, 655n, 772n  
 Thanissaro Bhikkhu 704-5  
 Thatcher, A. 568n  
 theology 12, 23-5, 158, 201-2, 337, 481-2, 484, 509, 513, 517-9, 523-5, 530-3, 535, 560, 563-4, 566-72, 654, 688, 733, 771n, 779  
 Theunissen, M. 519n  
 Thomas Aquinas, St. 24, 530-3, 535, 733, 755  
 Thompson, I.E. 568-9n  
 Thompson, M. 11, 181-93  
 Thoraval, J. 748n  
 Thrilling, L. 628-9  
 Tiantai Buddhism 27, 29, 604-6, 613, 664-5, 674-5, 677  
 Tiedemann, D. 779, 782  
 Tillich, P. 25, 563-4, 566-70, 572

- time 9, 19-21, 34, 42, 75, 77, 83, 88,  
92, 100-2, 105-7, 113, 118-128,  
130-3, 135, 139, 147, 152, 161-2,  
165, 174, 183-5, 188, 190-2,  
194-207, 211-2, 217, 220-3, 226,  
265, 287-90, 294, 296, 300, 324,  
329, 330, 373, 403, 406, 417,  
419, 426-8, 431-3, 470, 498, 501,  
506, 528, 541-2, 559, 567, 579,  
582, 647, 660, 687, 697, 707,  
715, 727, 743n, 744, 793, 795-6,  
798
- Timmerman, J. 5, 52n, 273n, 386n  
Timmons, M. 247-9, 255, 699n  
Trampota, A. 386n  
transcendental idealism 9, 16, 29, 31,  
62, 101n, 110-3, 116-7, 125,  
133-4n, 135, 138, 169, 201, 218,  
220, 223, 712, 727-9, 799, 802,  
806
- Tu, C.I. 748n  
Tu, W.M. 322n, 326n  
Tuck, A.P. 683  
Tuck, R. 392n, 393n  
Tugendhat, E. 653-4n  
Tuschling, B. 394n
- Unger, U. 67n  
Urmson, J.O. 106n
- Vaihinger, H. 475n, 655, 658  
Van Cleve, J. 136n, 138n  
Van Impe, S. 20, 424-37, 320n  
Van Norden, B.W. 326n  
Vandenabeele, B. 15-6, 308-20, 437n  
Vanhoozer, K. 571n  
Vasubandhu 30, 709-10, 712-4  
Vaughan, G. 397n  
Velleman, D. 666  
Vitoria 755, 758  
Voltaire 656n, 765  
Vorländer, K. 221n, 754  
Vos, R. 33, 764-76  
Vouvet, J. 657n
- Waley, A. 67  
Walker, R.C.S. 224  
Waller, B.N. 258n, 261n  
Walzer, M. 756  
Wang, B. 83  
Wang, Q. 592n  
Wang, R. 333n  
Wang, Y.M. 27, 613, 621-2, 631-2,  
636  
Ware, O. 699n, 702n  
Warner, M. 571n  
Warnock, G.J. 106n  
Warren, M.A. 252n  
Watkins, E. 219n, 228n  
Watson, B. 322n, 333n, 344n, 674n  
Watt, J. 673n  
Weischedel, W. 537n, 791n  
Weiss, R. 351n  
Wellenreuther, H. 783n  
Wenning, M. 28-9, 653-63  
Wenzel, C.H. 16, 321-32, 345n  
Weyland, K. 793n  
Wieland, W. 520n  
Wike, V.S. 642n  
Wilcke, J.C. 301  
Wilhelm, R. 419n, 769n, 813n, 820n  
Willaschek, M. 49  
Williams, B. 250  
Wils, J.P. 547n  
Wimmer, R. 547n, 548n  
Wittgenstein, L. 24, 65, 536-41, 544-  
9, 742  
Wittmann, R. 394n, 397n  
Wodarszik, U.F. 11, 167-80  
Wohlfart, G. 4-6, 16, 53-73, 342, 344  
Wolff, C. 33, 161, 279n, 281n, 335-6,  
765, 766, 768-71, 775, 777, 781  
Wolff, R.P. 129n  
Wong, K.K. 118-28  
Wood, A.W. 109-10, 254, 305n, 350,  
352-3, 355-7, 363, 393n, 429n,  
430n, 433n, 546n, 696n, 701,  
770-1  
Wright, R. 258n  
Wuerth, J. 10-1, 141n, 157-66
- Xiaojing (Book of Filial Piety)* 82  
Xie, S.S.J. 8, 32n, 34-5, 799-807  
Xie, S.X.L. 26-7, 592-602  
Xunzi 28, 80, 94, 639, 645-52

- Yang, G. 333n  
Yazdi, M. 735  
Yovel, Y. 486  
Yu, E. 381n  
Yu, H. 591n  
Yu, J. 360n
- Zammito, J.H. 311n, 342, 783-4  
Zanetti, V. 763n  
Zengzi 79, 82, 91  
Zhang, E. 29-30, 678-91  
Zhang-Kubin, S. 67n  
Zheng, X. 83  
Zhi Yi 604n
- Zhong Yong* 84-7  
Zhou, D. 53-4  
Zhu Xi 86-7  
Zhuangzi 6, 53, 66, 69-72, 333-4,  
337-8, 344n, 345-6, 631, 660-1  
Zielinski, S. 300n  
Zigong 83  
Ziporyn, B. 333n, 334n, 344n  
Zisi 80, 84, 87  
Zitelman, A. 753-4n, 757n  
Zixia 323, 325  
Zöller, G. 19, 386n, 401-14  
Zuckert, R. 308n