Kant's Ethics of Virtue

Edited by
Monika Betzler

DE GRUYTER

Kant's Ethics of Virtue

Kant's Ethics of Virtue

Edited by
Monika Betzler

Walter de Gruyter · Berlin · New York

∞ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

A CIP catalogue record for this book is available from the Library of Congress.

ISBN 978-3-11-017728-2

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.

Printed in Germany
Cover design: Christopher Schneider, Berlin.
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen.

Contents

Notes on Contributors

Elizabeth Anderson is John Rawls Collegiate Professor of Philosophy and Women's Studies at the University of Michigan, Ann Arbor. Her research focuses on democratic theory, egalitarianism, the ethical limits of markets, theories of value and rational choice, the philosophies of John Stuart Mill and John Dewey, and feminist epistemology and philosophy of science. She is the author of *Value in Ethics and Economics* (1993). She is currently writing a book on the ideal of ethno-racial integration in democratic theory.

Marcia Baron is Rudy Professor of Philosophy at Indiana University, Bloomington. Her main research areas are Kantian ethics and philosophical issues in criminal law. She is the author of *Kantian Ethics Almost without Apology* (1995), and she co-authored (with Philip Pettit and Michael Slote) *Three Methods of Ethics* (1997). She published numerous articles on topics in ethics, political philosophy and philosophy of law, including friendship and impartiality, patriotism, justifications and excuses, sexual consent, and manipulativeness.

Monika Betzler is Professor of Philosophy at the University of Berne. Her research encompasses theories of practical reasoning and rational agency, moral psychology, and ethics. She has written a book on self-knowledge, *Ich-Bilder und Bilder-Welt* (1995), and just completed a book manuscript, *Personal Projects as Practical Reasons.* She edited (with Barbara Guckes) a collection of articles on Harry Frankfurt's philosophy, *Autonomes Handeln* (2000), and (with Peter Baumann) *Practical Conflicts* (2004). She has published various articles on the normative role of emotions, the rational explanation of action, practical conflicts, and autonomy.

Stephen Darwall is John Dewey Distinguished University Professor of Philosophy at the University of Michigan. He has written widely on the foundations and history of ethics and is the author, most recently, of *The Second-Person Standpoint: Morality, Respect, and Accountability* (2006). His other books include: *Impartial Reason* (1983), *The British Moralists*

and the Internal Ought: 1640–1740 (1995), *Philosophical Ethics* (1998), and *Welfare and Rational Care* (2002).

Andrea Esser is Professor of Philosophy at the Philipps-University of Marburg. Her areas of research include ethics, philosophy of law, and aesthetics, with a focus on Kant and ancient philosophy. She is the author of a book on Kant's theory of art, *Kunst als Symbol. Die Struktur ästhetischer Reflexion in Kants Theorie des Schönen* (1997), and a book on Kant's doctrine of virtue, *Ethik für Endliche. Kants Tugendlehre in der Gegenwart* (2004). She is the editor of Kant's *Critique of Judgment* for the new Akademie edition of Kant's works, and has published a number of articles on Kant's moral philosophy, ethics, and aesthetics.

Thomas E. Hill, Jr. is Kenan Professor of Philosophy at the University of North Carolina, Chapel Hill. His essays on moral and political philosophy are collected in *Autonomy and Self-Respect* (1991), *Dignity and Practical Reason in Kant's Moral Theory* (1992), *Respect, Pluralism, and Justice Kantian Perspectives* (2000), and *Human Welfare and Moral Worth: Kantian Perspectives* (2002). With Arnulf Zweig he co-edited a new edition of Kant's *Groundwork for the Metaphysics of Morals* with extensive notes (2002).

Christoph Horn is Professor of Philosophy at the University of Bonn. He works on various aspects of ancient philosophy (Plato, Aristotle, Plotinus, and Augustine) as well as modern moral and political philosophy (especially on Kant). He is the author of *Plotin über Sein, Zahl und Einheit* (1995), *Augustinus* (1995), *Antike Lebenskunst* (1996), and *Politische Philosophie* (2003). He co-authored (with Corinna Mieth and Nico Scarano) a commentary on Kant's *Groundwork* (2007), and co-edited (with Dieter Schönecker) a volume of essays on Kant's *Groundwork* (2006).

Robert N. Johnson is Associate Professor of Philosophy at the University of Missouri, Columbia. His articles on virtue, practical reason and Kant's ethics have appeared in *Ethics*, *The Philosophical Quarterly, Philosophy and Phenomenological Research*, *Pacific Philosophical Quarterly* and *The History of Philosophy Quarterly.*

Samuel J. Kerstein is Associate Professor of Philosophy at the University of Maryland, College Park and Fellow in the Program in

Ethics and Health at Harvard University. He is the author of *Kant's Search for the Supreme Principle of Morality* (2002), as well as several articles on the foundations of Kantian ethics, including most recently "Deriving the Formula of Humanity" (in *Groundwork for the Metaphysics of Morals*, Christoph Horn/Dieter Schönecker, ed., 2006). He is currently working on an elaboration and defense of Kant's Formula of Humanity. Reflection on issues in bioethics such as the moral permissibility of purchasing organs for transplantation informs Kerstein's efforts to reconstruct this principle.

Jörg Schroth teaches philosophy at the Universities of Göttingen, Munich, and Berne. His research areas include ethical theory, Kant's practical philosophy, and the work of the Neo-Kantian German philosopher Leonard Nelson. He has written a book on the universalizability of moral judgment, *Die Universalisierbarkeit moralischer Urteile* (2001), and is currently editing Leonard Nelson's *Nachlass.* His articles on various topics in ethics have appeared in *Inquiry, The Journal of Value Inquiry, Logical Analysis and History of Philosophy,* and *Zeitschrift für philosophische Forschung.*

Philip Stratton-Lake is Professor of Philosophy at the University of Reading. His main area of research is moral philosophy, Kant and intuitionism. He is the author of *Kant, Duty and Moral Worth* (2000), and Editor of *Ethical Intuitionism: Re-Evaluations* (2002), W.D. Ross's *The Right and the Good* (2002), and *On What We Owe to Each Other* (2004). He has also written numerous articles on Kant, Ross, Intuitionism, and metaethics.

Katja Maria Vogt is Associate Professor of Philosophy at Columbia University, New York. She is the author of a book on Pyrrhonian Scepticism, *Skepsis und Lebenspraxis* (1998), as well as *Law, Reason, and the Cosmic City: Political Philosophy in the Early Stoa* (2007). She has published various articles on ancient philosophy and ethics.

List of Abbreviations

The following abbreviations for Kant's works are used throughout this book and will be followed by the volume and page number from the German acadamy edition (Ak.): Kant's Gesammelte Schriften, edited by the „Königlich Preussische Akademie der Wissenschaften" (Berlin: Georg Reimer, later Berlin: de Gruyter, 1902 ff.). References to the German acadamy edition will be in square brackets. References to other editions and English translations of Kant's works are listed by each contributor individually.

APPV Anthropology from a Pragmatic Point of View (Anthropologie in pragmatischer Hinsicht abgefasst)
CPrR Critique of Practical Reason (Kritik der praktischen Vernunft)
G Groundwork of the Metaphysics of Morals (Grundlegung zur Metaphysik der Sitten)
MM The Metaphysics of Morals (Die Metaphysik der Sitten)
R Religion within the Boundaries of Mere Reason (Die Religion innerhalb der Grenzen der blossen Vernunft)

Kant's Ethics of Virtue: An Introduction

Monika Betzler

Kant's moral theory is typically considered to be a prime example of an ethics of duty, sometimes categorized as deontological ethics. After all, what we are morally required to do, according to Kant, can be derived only from a non-empirical standard of rationality, namely the Categorical Imperative. It provides the sole principle for moral evaluation. It is authoritative precisely because the conceptual features of our willing yield what is right for us to do. And they provide that standard of rightness unconditionally, and thus independently of the empirically given inclinations and dispositions of the agent and the consequences of his actions. All duties are thus thought to be derived from reason, and their justification is independent of any empirical facts about a person's feelings or other characteristics.

The title of this anthology may thus seem misleading at best. To speak of "Kant's ethics of virtue" almost sounds like a contradiction in terms. What, if anything, could Kant share with virtue ethics? It is not simply that contemporary virtue ethicists more often than not conceptualize themselves as an antidote to Kantian ethics. What is more, virtue – either as a generalized disposition or as a set of particular traits of a person – simply does not figure as a core feature of his moral theory. And this is not surprising given Kant's central project of providing a non-empirical and therefore secure foundation for morality. This is why he focuses on an analysis of the rational will and a conception of self-governing reason, and not on what kind of person one should be.

In opposition to the Kantian project, virtue ethics focuses on character, or what kind of person one should be. What turns virtue ethics into a distinct normative theory that differs from deontology and consequentialism alike is the idea of human flourishing as an activity, or more generally, as an ideal of how to live well. Virtues are then derived from natural conditions of mankind: from facts about what makes us flourish. They are determined by what is regarded as the natural end and natural powers of human beings. What is foundational to virtue ethics, then, is not right or duty, nor value or the good, but an ideal of the

person or what makes for a life well lived. Accordingly, to find out what actions are right, and what things or states of affairs are good, one has to resort to an ideal of the person. The standard of right action is to do whatever an entirely virtuous person would do. And the particular view of the ideal of the person must be independent of any understanding of the right or the good.

The contrast just canvassed between Kant's moral theory and virtue ethics is not altogether wrong, but it would be hasty to conclude that virtue does not play an important role in Kant's theorizing. The aim of this collection is to shed new light on the connection between Kant's moral theory and the role played by virtue in particular and the themes of virtue ethics more generally. To foster debate on how this connection is to be conceived is no mere academic enterprise. Kant himself, and particularly the views expressed in his later ethical writings, render support to raising questions about this connection.

Kant's *Groundwork of the Metaphysics of Morals* (1785) and his *Critique of Practical Reason* (1789) are generally taken to be the major sources for his ethical theory, and they dominate the assessment of Kant's views. Kant's later writings have not only received less scholarly attention than his earlier works,[1] but what frequently remains unnoticed – or, to put it more cautiously, what is noticed too little – is that Kant developed, extended, and sometimes modified his views in his later writings. Most notably, the second part of his *Metaphysics of Morals* (1797), the "Doctrine of Virtue", provides ample new material, particularly with regard to the ideal, and thus virtuous, moral agent. In addition, his *Anthropology from a Pragmatic Point of View* (1798) and his *Religion Within the Boundaries of Mere Reason* (1793) enrich his earlier insights. Kant went in many ways beyond what is traditionally viewed (and often criticized, if not caricatured) as the core of his ethics as he developed it in his *Groundwork* and *The Critique of Practical Reason.* I will highlight these new themes in due course.

Among the topics that figure most in Kant's later writings and that are taken up in the contributions to this anthology, let me just mention here the following: In the *Metaphysics of Morals* Kant provides a

1 This has been remedied only somewhat recently. See, for example, Mark Timmons (ed.), 2002, *Kant's Metaphysics of Morals.* Interpretative Essays, Oxford: Oxford University Press; and Andrea Esser, 2004, *Eine Ethik für Endliche.* Kants Tugendlehre in der Gegenwart, Stuttgart: Frommann-Holzboog.

conception of moral excellence and the ideal moral person by introducing duties of virtue. He draws attention to two ends that he considers obligatory, namely, one's self-perfection and the happiness of others. It is a duty of each rational agent to pursue these ends, but this duty can be performed with more or less excellence. Kant famously distinguishes between perfect and imperfect duties and elaborates on how one's duty to perfect oneself and to realize others' happiness are imperfect: We have the duty to adopt some end that instantiates these general ends, but there is leeway as to which specific course of action we take. His notion of imperfect duty provides theoretical resources for tackling practical conflicts between different duties. Also, he emphasizes much more in his later work what dispositions an ideal moral agent should develop. He introduces his understanding of virtue as strength of will, and concentrates on moral emotions and feelings, such as love and respect, and the role they play for us in doing our duty.

The connection between Kant's moral theory and virtue is thus established by Kant himself. And it is this connection that helps us understand Kant's own development, and to elucidate later themes that provide important answers to morally pertinent questions as they are raised by current debates within ethical theory. Also, as will become clear in the contributions to this collection, Kant's later ethical writings provide an excellent source for confronting objections to his ethics that have been raised by contemporary virtue ethicists. It is thus the purpose of this collection to further debate on themes that Kant himself developed and deepened only in his later ethical writings.

This volume brings together eleven essays by Kant scholars as well as moral theorists who find themselves inspired by Kantian ideas. Some of the chapters focus more on the development of Kant's ideas from his earlier to his later work and attempt to make better sense of some core theoretical elements of his moral theory. Others start from an independent interest in a particular morally pertinent question and then develop an answer in light of Kant's later insights. In what follows, I will try to delineate the ways in which Kant's moral theory relates to themes of virtue ethics, and what questions and topics the essays in this collection bring to the fore. In common with the contributors, I presuppose some familiarity with the basic insights of Kant's moral theory.

The contributions to this volume can be grouped around four different yet overlapping themes. One theme is the role that virtue really plays in Kant's work and how it relates to his ethics of duty (Hill, Johnson, Schroth, Stratton-Lake). These contributions help to clearly mark the distinction between Kant's deontological theory and virtue ethics in its ancient or modern form. Even though Kant puts forth a "Doctrine of Virtue", and the somewhat provocative title of this collection seems to suggest such a reading, Kant's ethics of virtue is clearly distinct from classical or contemporary virtue ethics. But this does not preclude the possibility that virtue plays a more important role in Kant's later work.

Another theme is the importance of the ideal agent's motivation, particularly as it is reflected in his emotional attitudes (Anderson, Horn). A third theme concentrates on the idea of persons as ends in themselves, to be respected and not merely treated as means (Darwall, Kerstein). And a final theme more directly confronts criticism raised against Kant's ethics by providing a more thorough analysis of his contrast between perfect and imperfect duties. Kant's conception of imperfect duties helps to show that his moral theory is not too demanding, in that duties to others are limited by duties to develop one's own non-moral capacities. In a similar vein, and contrary to criticisms that Kant's ethics is overly impartialist, it is Kant's conception of imperfect duties that provides the room for love and friendship. Finally, imperfect duties also provide the means to rationally resolve practical conflicts. They thus help to forestall the worry that Kant's ethics is overly rigoristic (Vogt, Baron, Esser).

Virtue, the Good, and the Right

Kant considers virtue to be an ideal with respect to the source of one's actions. He regards virtue as "fortitude with respect to what opposes the moral disposition within us" [*MM* 6: 380], i.e., strength in adhering to duty in the face of one's inclinations, or "the moral capacity to constrain oneself" [*MM* 6: 394]. Hence, in contrast to most moral theorists, Kant regards virtue as a kind of strength of will to do what is right. Virtue must result from considered principles [cf. *MM* 6: 384–5], and to be virtuous is, most centrally, to act from duty.

Thomas Hill discusses Kant's conception of moral virtue and its place in his ethical theory in his "Kant on Virtue and 'Virtue Ethics'". He sets the stage for an assessment of the relation between what is generally thought to be Kant's ethics of duty and virtue ethics. According to Hill,

Kant's claim that an act has moral worth only if it is done from duty should not be viewed as a complete theory about how to assess the moral value of actions. It is true that the fact that something is morally required should provide sufficient reason for doing it. But this does not mean that an agent must always be thinking of duty. Hill distinguishes between two perspectives from which the moral value of acts can be evaluated. One perspective is the agent's will to act as he should, even if he is inclined otherwise. The other perspective is the agent's success in acquiring character traits that help to live a moral life. These character traits should be cultivated even if success in doing so is not entirely up to the agent.

Hill also provides a reading of Kant's account of virtue as strength of will. A good will as a fundamental commitment to doing what is right may also be weak. A virtuous person's will is both good and strong. To understand how virtue is more than a good will, Hill discusses what Kant could mean by moral weakness. Morally weak persons have a good will, but on particular occasions they fail to live up to their basic commitment to do what is right. Hill takes the morally weak to behave according to two conflicting maxims: a basic maxim of conforming to morality's unconditional requirements, and a shorter-term maxim of putting self-love above morality on a particular occasion. Morally weak persons need to make better use of their capacities to adopt and implement a commitment to do what is right.

Hill also examines whether an expanded Kantian ethics might accommodate concerns that virtue ethicists could raise against it. He spells out in what ways Kant's ethics is different, but also shows how virtue ethical concerns might be alleviated. First, it may be questioned whether moral rules or principles are necessary for individual decisions. Virtue ethicists might think instead that it is better to promote the development of one's character. Kant, to be sure, proposes principles as necessary for a morally good life. He thinks that everyone needs to respect such principles in order to be fully virtuous. Kant differs from virtue ethics in that he considers the development of character to be not sufficient, but takes judgments of practical reason as essential. According to a second charge of virtue ethicists, any moral rule is liable to oversimplify the complexities of a particular situation. But what speaks for Kant's view is that, at some level, there must be some common ground for the legitimate variation of particular moral judgments. Third, it might be maintained that Kant ignores relevant features of good character and motivation. It is true that Kant emphasized that emotions do not serve as grounds for moral judgment. He also frequently pointed out that

emotions are likely to distort our moral judgments. But Kant did acknowledge some role for emotions, such as respect, pleasure at having done what is right and guilt for having knowingly done wrong. Kant also pointed out that positive other-regarding feelings should be cultivated and make it easier for us to do our duty. Such emotions, however, are the product of our moral awareness, not its source. Fourth, concerning the standard of right action, virtue ethicists subscribe to the view that the right act is the act that a virtuous person would do in the situation. They then specify what a virtuous person would do. Hill points out that to the extent that virtuous character is defined in part as respecting the principles of justice, honesty, friendship, and so on, the difference to Kantianism is minimal. They would share the view that there are independent standards of right action. If extreme virtue ethicists, however, claim that there are no independent standards for right action, they substantially differ from Kant's view. Kant would regard such an idea of virtue to be too insubstantial to guide our moral decisions. What is more, we must already have a standard of right in order to identify a person as an exemplar of virtuous character. To be virtuous thus is in part to know and respect independent reasons for acting rightly. We should do what a virtuous person would, but when it comes to the justification of morality, our standards of the right and the good cannot be derived from the idea of virtue.

Robert N. Johnson also examines what Kant's ethics might share with virtue ethics. According to Johnson, it is Kant's notion of the good will that may be thought to offer the "form" that supports the view that these two ethical doctrines have more in common than one may think. After all, the notion of the good will could characterize an ideal of the person for rational agents, and thus serve as a foundation to generate a rule of dutiful action, namely the Categorical Imperative. To play that role, the good will would have to be regarded as a final end that provides the reasons to guide deliberation. Kant thus may be thought to ground the Categorical Imperative in a final end, the good will and thus in the humanity of persons. The difference between virtue ethics and Kant's ethics then concerns only the content of an ideal for persons. What they share is the "form" in that both take the ideal of a person as an end.

Moreover, Johnson shares Hill's diagnosis that one of the advantages of Kant's ethics is that the good will is not equivalent to a virtuous will. The ideal person is therefore taken to be someone with non-ideal

character traits. He is thus able to explain why we should improve our characters.

But, as Johnson points out, even though Kant's ethics has the same form as virtue ethics in providing an ideal of the person, it differs quite fundamentally in that it does not assign the ideal of the person the same normative priority. The good will cannot be understood independently of the Categorical Imperative. What renders a good will good is the disposition to deliberate and act by way of respecting the authority of the Categorical Imperative. The good will thus is not the fundamental notion out of which a theory of right action can be generated. It is because the Categorical Imperative binds the will of an agent that the good will (and humanity) have value. Against virtue ethical readings of Kant, his notion of the good will cannot be thought to hold the same position as human flourishing in virtue ethics. It is not the case that Kant's concept of duty is based on a conception of value (namely, on his notion of the good will). Similarly, Kant's conception of rational agents is informed by a prior understanding of the moral law. They can therefore not provide a rationale for the moral law. Hence, even if Kant's ethics is taken to introduce an ideal of the person, and thus shares what Johnston calls the "form" of virtue ethics, it differs from substantiality in that morality is justified independently from such an ideal.

In highlighting and elaborating what Kant's moral theory, and some of the core features of the second part of his *Metaphysics of Morals* in particular, shares with virtue ethical themes, both Thomas Hill and Robert Johnson subscribe to the view that Kant's ethics differs in important respects from virtue ethics. While Hill emphasizes how Kant would respond to virtue ethical concerns raised against his theory, Johnson adds a more sceptical analysis, and thus emphasizes in which respect both theories remain distinct, most notably with regard to the foundation of morality.

This assessment is echoed by Jörg Schroth's more systematic analysis with regard to the relation of the right and the good in normative ethics. It is in the "Doctrine of Virtue" that Kant grounds our duties in the promotion of ends. This has led commentators to maintain that Kant accords the good (at least some) priority over the right.

Against such a reading Schroth defends the view in "The Priority of the Right in Kant's Ethics" that Kant continues to subscribe to the priority of the right. To bolster this claim Schroth first examines how the priority relation between the good and the right can be conceived at all.

He distinguishes between three kinds of relations: (i) the definition of the good (right) before the definition of the right (good), (ii) the definition of the good (right) independently of the right (good), and (iii) the non-existence of impermissible conceptions of the good. Schroth shows how none of these conceptions can give a plausible meaning to the priority of the right or good. They cannot serve to distinguish deontology and Kant's ethics from consequentialism, and do not in fact correspond to Kant's own view on the priority of the right.

He discusses an interpretation of the priority of the right in Kant's ethics according to which the right is defined without any reference to or presupposing anything about the good. An action would thus be viewed as right under any circumstances, no matter what the consequences. Schroth argues that such a view would rob the priority of the right of any significance and would distort Kant's own view. This distorted view is attributed to a widely held misunderstanding of consequentialism according to which the right has merely instrumental value when it demands the maximization of the good. But, as Schroth shows, the demand for the maximization of the good is compatible with an autonomous and non-instrumental conception of the right. Consequentialism is therefore not tied to the view that the right is a derivative concept. This is why the difference between consequentialism and Kant is not one between the priority of the good and the priority of the right. Moreover, one does not need to claim that the right is completely independent of the good in order to secure the non-instrumental value of the right. Instead, consequentialism too has an autonomous conception of the right: The maximization of the good thus functions as a categorical imperative.

Moreover, Kant also has a conception of the good that is independent of the outcome of a right action. It is this independent conception of the non-moral good that allows Kant to distinguish between states of affairs that are good and those which are not good, even if both result from a good will.

As Schroth argues, the priority of the right does not distinguish Kant's ethics from consequentialism at all. Both theories employ the same concepts of the right and the good, and both theories subscribe to the priority of the right in its only tenable sense, which is Kant's original meaning of the priority of the right, namely, that the right rather than the good is the determining ground of the will. Schroth concludes that in this sense the right has priority over the good in every moral theory, except virtue ethics. But it is precisely the priority of the right so understood that

starkly distinguishes Kant's ethical theory from virtue ethics. Thus, from a systematic analysis of the relation between the right and the good, Schroth lends support to Johnson's argument.

In his "Being Virtuous and the Virtues: Two Aspects of Kant's Doctrine of Virtue", Philip Stratton-Lake examines how, in Kant's moral theory, being a virtuous person is related to having specific virtues. He provides an in-depth analysis of Kant's own development from the *Groundwork* to the *Metaphysics of Morals* with regard to Kant's conception of being virtuous and having virtue. In the *Groundwork* Kant's account of morally worthy acts serves as the basis of his explanation of what it is to be a virtuous person. It is the general disposition to act from duty that appears to render a person virtuous. In his *Metaphysics of Morals* Kant develops an account of virtue from his earlier account of morally worthy actions. But he also seems to reverse the order. While morally worthy actions seem primary in the *Groundwork,* and virtue is defined with reference to them, Kant now defines virtuous action with reference to the agent's virtue, namely, as an action that realizes the agent's virtue as the disposition to act from duty. Stratton-Lake proposes to render the later version compatible with the earlier one by elucidating the difference between a virtuous act and a morally worthy act. He shows that acting from duty is only a necessary, not a sufficient, condition of virtuous acts. That is, one can act from duty even if one is not virtuous. He thus shares Hill's and Johnson's view, but also adds a further interpretation of virtuous action: Virtuous actions are done from duty, but they must also realize a general disposition to fulfil all of one's obligations from duty. A virtuous person has such a general disposition. As a result, a specific disposition to act from duty in specific circumstances has to be distinguished from the general disposition to act from duty.

There is one general virtuous disposition, but there are several virtues corresponding to the several duties of virtue. Hence, being a virtuous person is contrasted to the plurality of particular virtues. Conformity with the various duties of virtue determines these virtues. To make sense of what these duties of virtue require, we need to understand what Kant means by imperfect duties. For that purpose Stratton-Lake suggests distinguishing between their deontic valence and their deontic force. Whether some act is thought to be required or wrong constitutes its deontic valence. While perfect duties proscribe which acts we must not do, imperfect duties tell us which of the acts that are not wrong are required.

As far as the deontic force is concerned, perfect duties require us not to do certain acts, while imperfect duties require us to adopt certain ends (and hence the relevant maxim of ends) and thus allow for some latitude in complying with them. In the *Groundwork* Kant appeared to hold that we do not have to act according to our ends if we are not inclined to do so. This suggests that imperfect duties have less deontic force. His later view, developed in the *Metaphysics of Morals,* however, is more plausible, albeit more demanding: We have some latitude in realizing an end, although not in the sense that we can trade it against the interest of our inclination; we can trade it only against some other imperfect duty. The difference between perfect and imperfect duties seems better captured by the difference in deontic content (rather than in deontic force): Perfect duties prohibit and imperfect duties require with the same force, but they require different things.

To be virtuous is not simply to comply with all duties of virtue, but to hold the general disposition to comply from the motive of duty. To have some specific virtue is for Kant to act in accordance with the relevant duty of virtue and comply with it from the motive of duty. Compliance from the motive of duty is a necessary condition for a specific virtue.

Stratton-Lake points to a major flaw in Kant's "Doctrine of Virtue": Virtue or the state of being virtuous cannot be realized in the exercise of the specific virtues, since virtuous acts are simply those that realize the agent's disposition to act from duty. Since there is no duty to beneficent acts, for example, one cannot knowingly perform such acts from duty. Stratton-Lake considers this to be "a very odd consequence of any theory of virtue". And even if beneficent acts were morally required, another feature of imperfect duties would be lost, namely, latitude. As a result, Stratton-Lake proposes to revise Kant's theory by introducing so-called disjunctive duties. They are duties to do either A, or B, or C. One of several actions can thus be done from duty, even though no single action is obligatory. It is the disjunction that is obligatory. Disjunctive duties thus allow for latitude. Actions that realize specific virtues could thus be done from duty and realize the agent's virtue as such. In addition, disjunctive duties have the advantage of placing plausible limits on the kind of latitude because implausible options are not taken to fare as disjuncts of a disjunctive duty. Stratton-Lake claims that incorporating disjunctive duties into Kant's account of virtue yields a compelling alternative to Aristotelian-inspired accounts of virtue.

The Emotions: Respect, Self-Esteem, and Love

Kant is famous for denying that emotions, feelings, and other inclinations offer a suitable foundation for morality. Only actions done from duty have moral worth. It is not only that some of our emotions and feelings are not directed to any moral ends. What is more, even moral emotions cannot be summoned at will; they are still likely to be expressive of our self-love, and thus cannot give rise to obligation. Thus, actions done from (moral) emotions do not deserve any moral merit.

In her "Emotions in Kant's Later Moral Philosophy: Honour and the Phenomenology of Moral Value", Elizabeth Anderson shows that Kant was more sympathetic to emotions and feelings than his remarks on morally worthy actions in the *Groundwork* suggest. His "Doctrine of Virtue" calls for the cultivation of many emotions, and reflects a much richer phenomenology of feelings than his earlier work.

According to Anderson's interpretation, Kant draws a distinction between two different forms of value that she takes to be central to his moral theory. The difference between these different forms of value can be grasped through our feelings: There are what she calls "appealing values" that constitute the domain of the good, and there are what she calls "commanding values" that belong to the domain of the right. While appealing values are ends we adopt according to standards that are up to us, commanding values themselves carry authority and represent a constraining standard for us. While appealing values tempt us, commanding values necessitate. And it is this distinction that seems to be reflected in the vast variety of our feelings. Respect for the moral law and its source is a feeling that reflects a commanding value. But negative feelings of shame, guilt, and contempt, and positive feelings of contentment, pride, and self-esteem, are also responses to a commanding value. The Kantian agent is thus moved by a set of feelings toward himself with respect to how he fails in or lives up to his duty. Without such feelings, agents would not feel moral demands to be compelling. Feelings of appeal, by contrast, are captured by affects, wishes, desires, passions, and aesthetic taste. With regard to morality, they can play an important instrumental role in helping us do our duty. Similarly, they can impede the fulfilment of our duty.

According to Anderson's reading, the fundamental Kantian moral feelings are respect and self-esteem. In locating their historical source in the ethics of honour, she elucidates how commanding value is connected to the value of persons. According to an ethics of honour, persons are

deemed worthy because they occupy certain positions in a hierarchical society, and they have a sense of self-worth because of the commanding value they embody. In taking persons to deserve respect not for their standing in society but for their standing as rational agents, Kant transformed the ethics of honour to an ethics of respect. He thus reverses the relations of respect and esteem, a development Anderson considers to be "revolutionary". While in the ethics of honour persons deserve respect depending on the esteem of others, in Kantian ethics persons deserve moral esteem by virtue of their respect for the other's humanity as one's end.

Anderson hastens to point out, however, that Kant does not always manage to leave the ethics of honour behind, and thus fails to formulate an entirely consistent morality of universal respect. This is revealed in his argument that rape victims behave dishonourably in succumbing to rape rather than resisting to their deaths. In a consistent ethics of respect, honour must be dependent on respect for humanity, and hence on the duty to protect human life. Rape victims should therefore not give up their right to life for the sake of honour.

Another topic which becomes prominent only in Kant's later writings on ethics is the moral significance of love. In the "Doctrine of Virtue" Kant divides all "duties of virtue" into those toward human beings and those toward non-human beings. The duties of virtue owed to human beings are then further specified as those we have toward ourselves and those we have toward others [*MM* 6: 413]. Duties to others are subdivided into "duties of love" and "duties of respect" [*MM* 6: 448].

In his earlier work Kant emphasizes that love as an emotional inclination cannot be commanded. What he calls "practical" and not "pathological" love is an attitude that moves agents to practice beneficence. It is a moral duty that obligates independently of any emotion. Love so conceived seems to be an equivalent of the Categorical Imperative. By contrast, the "duties of love" that Kant introduces in his "Doctrine of Virtue" seem to be much more restricted in scope: Each agent is thought to have a more specific duty to adopt the maxim to practice beneficence, gratitude, and sympathy.

In "The Concept of Love in Kant's Virtue Ethics", Christoph Horn tackles what he calls the "problem of the wide or narrow application of love". According to Horn Kant introduces a much more precise account of love in the "Doctrine of Virtue". This account sheds new light on Kant's earlier remarks and helps to show that it is wrong to suppose (as

the earlier writings seem to suggest) that Kant distinguishes between an emotional and a non-emotional kind of love, that non-emotional love is a kind of attitude that can be commanded, and that he equates non-emotional love with beneficence. It is not the case that there is a dissonance between Kant's earlier and later works, but his later and more extended remarks on love show somewhat surprisingly that all types of love are emotional, that no kind of love can be commanded, and therefore that love is morally inappropriate. What morality demands is not love, but beneficence. To the extent that love can arise from practicing beneficence, love can be appropriate as the habituated consequence of a moral attitude.

As Horn's reading of the "Doctrine of Virtue" reveals, there is no wide or narrow application of love. Instead, Kant introduces two specific and non-conflicting aspects of love in his earlier and later work respectively. The allegedly comprehensive use of love is meant to capture the emotional state of an ideally moral person. She would not only practice beneficence, but would also have obtained a state of benevolence. The restricted use is concerned with the maxim to practice beneficence, gratitude, and sympathy. In his "Doctrine of Virtue" Kant focuses on ends that are moral duties, whereas in his earlier work he started with maxims quite independent of any ends. To the extent, however, that one starts with morally transformed ends, namely, one's own perfection and other people's happiness, we have to have maxims to achieve those ends. Benevolent love as well as respect are thus means to achieve the happiness of others. But Kant admits only an indirect role for love and other moral feelings because they have value only to the extent that they are derived from practical reason. Horn thus shows how Kant assigns love a more important role in his later work without giving up the idea developed earlier that love is derived from practical reason.

The Duty of Respect and Persons as Ends in Themselves

Both Horn and Anderson concentrate on emotions in their role of helping us do our moral duty, which Kant came to work out in his later writings. To the extent that Anderson highlights the importance of respect, she does so to emphasize how it reflects, as a particular emotion, the commanding value of the moral law as it is embodied in persons. Stephen Darwall, by contrast, looks at respect from a slightly different

angle. He shows how the duty of respect is directed to treating persons as ends in themselves.

It is a familiar Kantian doctrine that all rational beings have a dignity that gives them an equal claim to respect. In his "Kant on Respect, Dignity, and the Duty of Respect", Darwall canvasses the development of Kant's doctrine of respect: In the *Groundwork* and in the *Critique of Practical Reason* Kant somewhat surprisingly, and contrary to what the standard reading suggests, characterizes the dignity of persons in a more complex and puzzling way. In his earlier work, Kant conceives the dignity of persons in part, at least, as a form of merit. Various passages in the *Critique of Practical Reason* as well as in the *Groundwork* suggest that what we respect in the person of good will is her adherence to the moral law, and this seems to be what her merit consists in. If a person is to be an end in herself, she has to fulfil a certain condition, namely having morally worthy maxims. This is why a person's dignity is identified with her moral worth, and thus merit.

The "Doctrine of Virtue" is the first place where it becomes unambiguously clear that dignity warrants respect regardless of any form of merit (moral or otherwise). In addition, all rational persons have an equal standing in the demand for respect from one another. While respect for merit is what Darwall elsewhere called "appraisal respect", "recognition respect" is realized in recognizing its object's authority to make claims and demands of one another to respect, and therefore shows itself in how we act in relation to the object. In the "Doctrine of Virtue" Kant adds an interpersonal, or what Darwall calls "second-personal" dimension, according to which all persons have a unique authority to make claims and demands on one another's respect, and so hold one another responsible as equals.

This authority calls for recognition and is thus irreducibly second-personal. It calls for respect for the authority to make that claim. Hence, Darwall takes this idea of second-personal authority to be central to the dignity of persons. It requires that, if a person accepts a claim to respect, she must also accept that others have grounds for complaint if their claim to respect is violated. That is, making a valid claim or demand always presupposes the authority to make it, and the duly authorized claim creates a reason for compliance and responsibility.

What Kant must mean by the "duty of respect" in his later work is, according to Darwall, that we must limit our claims to authority by the equal (second-personal) authority that any person has: "Respect for oneself as a person thus involves insisting on one's equal second-personal

authority to make claims and demands *as* a person, where part of what one can demand of oneself and others is *respect for this very (equal second-personal) authority."* The duty of respect is a duty to have a specific *maxim,* the maxim of treating all persons as ends in themselves regardless of their merit, thereby according one another the authority to make demands on each other as equal persons. This idea of second-personal authority makes mutual accountability an important aspect of moral obligation. Darwall takes this to be the lesson of Kant's "Doctrine of Virtue", that compliance with the moral law and respect for the dignity of persons can only be established within the second-person standpoint.[2]

Darwall and Anderson thus share the view that Kant's notion of respect (particularly in his later writings) is that it is independent of a person's standing in society, and thus of merit. Both highlight how persons, according to Kant, deserve respect only for their standing as rational agents who make claims and demands on each other as equals.

While Darwall works out Kant's doctrine of respect as a duty to treat persons as ends in themselves, Samuel Kerstein approaches the question of what it means to treat persons as ends in themselves by way of an interpretation of what it means not to treat oneself merely as a means.

Kerstein sets out to elucidate Kant's rationale for perfect duties to oneself. Kant derives these duties to oneself from the Formula of Humanity: An agent should act so that a human being, whether in his own person or in the person of any other, is always at the same time treated as an end, never merely as a means [cf. *G* 6: 429]. To the extent that an agent acts contrary to these duties, he treats himself merely as a means. In his "Treating Oneself Merely as a Means", it is Kerstein's declared goal to elucidate what Kant could mean by treating oneself merely as a means. So far, Kant scholars have paid much more attention to the meaning of treating others as means. According to a received interpretation, an agent treats another merely as a means if he does something to which the other cannot consent. But this interpretation is not helpful in understanding what it could mean to treat oneself merely as a means. After all, killing oneself, masturbating, and lying – ways to treat oneself merely as a means, according to Kant – could be consented to and yet illustrate what it means to treat oneself merely as a means.

2 He has shown in his recent work how this can be done in light of Kant's insights. Cf. S. Darwall, 2006, *The Second-Person Standpoint*, Cambridge, Mass.: Harvard University Press, chs. 10–11.

The answer that Kant gives in the "Doctrine of Virtue" is the following: Treating oneself as a means amounts to failing to treat oneself as something that is unconditionally and incomparably valuable. But this answer, according to Kerstein, leaves mysterious how Kant arrives at some of our duties to ourselves. A lying agent, for example, may treat himself as a means yet he does not necessarily disrespect his value. Kerstein offers a different interpretation that stems from the duty not to make false promises: If an agent cannot rationally achieve the end he is pursuing in treating himself in a certain way, then the agent treats himself merely as a means. An agent cannot have this end just in case he cannot pursue it without thwarting the pursuit of some end that he is rationally compelled to have. Kerstein shows how this interpretation of treating oneself merely as a means sheds light on Kant's efforts to derive duties not to murder oneself, not to masturbate, and not to lie. In killing himself, an agent would render himself unable to attain an end that he is rationally compelled to pursue, namely that of refraining from destroying humanity. According to medical authorities of Kant's time, masturbation had deleterious effects on health, one of which was an inability to focus on any project. Accordingly, an agent would damage his capacity to pursue ends, a capacity that is central to his rational nature. Finally, in lying, an agent thwarts his attainment of something that he is rationally compelled to will, namely the accurate communication of one's thoughts. Kerstein concedes, however, that Kant does not offer a compelling justification for some of these derivations. Yet he takes this to be a flaw in the application of the Formula of Humanity, not a shortcoming of the formula itself. Kerstein thus provides an interpretation of perfect duties to ourselves as they are particularly worked out in Kant's later work.

The Demanding Nature of Morality, Impartiality, and Hard Choices: The Role of Imperfect Duties

Various objections have been lodged against Kantian ethics, particularly by critics inspired by virtue ethics. According to one important criticism, Kant's ethics is thought to be overly demanding, requiring us to constantly do our duty (typically understood as our duty to others).

In addition, it seems that practical reason requires us to entertain an agent-neutral or impartial perspective, testing maxims that are based on our personal and empirically given desires for their universalizability. We thus seem to be required to constantly act out of duty as Kant

understands it. The worry looming here is that we are overburdened with a concern for our adherence to duty. One may wonder, then, whether this makes the requirements of morality too demanding, and leaves no room for love and friendship. But while Kant does not seem to share this worry, he is quite aware of the limits of duties to others. Similarly, his conception of imperfect duties helps to make room for our partial bonds.

In the "Doctrine of Virtue" Kant spells out in more detail that practical reason prescribes moral demands of different kinds and that these demands form a system of duties. In her "Duties to Others: Demands and Limits", Katja Vogt suggests an interpretation of Kant according to which there are duties that limit our duties to others within a virtuous life.

One need not resort to supererogation in order to appreciate the limits of moral demands. Supererogationists hold (i) that the demands of morality are limited; (ii) that one can go beyond these limits; and (iii) that it is especially admirable to do so.

The idea that the demands of various kinds of duties are limited is grounded in Kant's distinction between perfect and imperfect duties, the latter allowing for latitude. The virtuous agent will thus figure out how to comply with all of her imperfect duties.

However, and as Vogt points out, while the duty to moral self-perfection, duties of respect, and duties of right shape and structure duties to others, they ultimately seem to make them more, or at least no less, demanding. Vogt shows that Kant cannot make sense of the key intuition of the supererogationist, who claims that there are actions which are not required but are good or virtuous. According to Kant no action can be considered good unless it is done from duty. As far as duties to others are concerned, any act of beneficence will be an act of duty. No matter how much effort an agent makes, her action will not go beyond duty. This feature of Kant's theory makes duties to others more rather than less demanding. It is part of our duty to others that we see acts as owed which are in fact not owed so as to spare others the feeling of humiliation.

But despite this feature, the duties to develop one's own non-moral capacities eventually limit the demands of duties to others in significant ways. This is because, according to Kant, the agent is closest to herself, and knows her "true needs" best. Vogt concedes that Kant does not offer a real criterion for limiting the demands of duties to others. But it seems that others are to keep a distance from the agent if they are to respect her.

And this leaves the agent room to figure out what her ends are and what is important to her happiness. Hence, the duty to self-perfection as well as the duty *of* others to respect us seem to limit our duties *to* others.

Marcia Baron attends to the objection that Kantian ethics is overly impartialist, or overvalues impartiality. She considers the "one thought too many" objection, according to which Kantians cannot allow an agent to take the fact that someone is, for example, his good friend or child as a genuine reason for doing something for that person. Instead, the objection goes, Kantians always need to resort to a deliberate assessment of their reasons from an impartial perspective, and this is precisely "one thought too many". Kant's ethics thus cannot leave room for love and friendship because an impartialist assessment of one's reasons renders it impossible to act out of love and friendship for a particular person. In her "Virtue Ethics, Kantian Ethics, and the 'One Thought Too Many' Objection", Baron argues that there is far less in Kantian impartialism for partialists to disagree with. Kant's conception of imperfect duties as he elaborates them in his "Doctrine of Virtue" allows for the latitude required to treat one's friends or relatives preferentially, or to give more due to one's subjective preferences. The duties to promote others' happiness and to seek one's own perfection leave it up to the agent's good judgment in deciding whom and how to help, or how to perfect oneself. To the extent that duties to particular others or the needs of others are not neglected or ignored altogether, it is permissible to attend more to one's friends and subjective preferences. As a result, partiality remains permissible for Kantians within moral bounds. The criticism of Kantians seems to confuse two different issues: (i) the fact that a person is, say, one's child does not render an immoral action morally permissible; and (ii) the fact that someone is one's child could never be relevant to the question of whether an action is morally permissible. Kantians do not disagree with (ii). Contrary to consequentialism, which may be the more appropriate target of the "one thought too many" objection, Kantian ethics has no requirement that we maximize impersonal value. This is why Kant's positive duties do not preclude meaningful relationships. As Baron points out, there is confusion concerning the level at which an impartial standpoint is called for. Whereas Kantians hold that actions must be justified from an impartial standpoint, they are not committed to the view that one should always entertain an impartial view. At the level of reasoning about what is permitted, we have to be strictly

impartial. But this does not mean that the fact that someone is my child, for example, is never morally relevant.

Finally, Baron points out that there is disagreement about how much we should take up an impartial perspective toward those we love or have relationships with. Even though she concedes that this is a genuine disagreement, Baron does not think that Kantians entertain one thought too many. Instead, Kantian ethics helps us to assess the moral claim that competing considerations make on us, and this can be done without undermining our love and attachments.

A further criticism that besets Kantian ethics concerns its rigorism as it reveals itself in the claim that the Categorical Imperative provides us with moral guidance for all possible cases. Moral conflicts thus do not seem to pose any problem or do not even seem to arise: The Categorical Imperative either proves one obligation not to be binding or simply requires us to follow the stronger one. But this seems to be quite at odds with our moral experience. More often than not circumstances are such that it is hard to come up with a clear moral judgment. Moral conflicts are pervasive and the fact that we are often tormented by regret after the rational resolution of such conflicts shows, according to critics, that Kant's principle-based moral theory is flawed. Contrary to Kant's claim, we sometimes seem to face conflicting normative demands. Since we can act on only one, we are forced to do wrong by breaching the other normative demand. This has led some critics to suggest that an ethics based on a single principle has to be abandoned.

Against this criticism Andrea Esser shows in her "Kant on Solving Moral Conflicts" how Kant provides the resources in his "Doctrine of Virtue" to tackle moral conflicts more adequately. According to Esser, Kant's ethics must not be misunderstood as a theory that is rigidly modelled on abstract principles, reducing moral evaluation to a subsumptive mechanism. Instead, moral deliberation is governed by formal principles, but remains open with respect to the concrete action taken.

As becomes evident in his later work, Kant does not regard the determination of the will as detachable from its realization. If the Categorical Imperative calls for action, however, it does not choose the action itself but its underlying determination of choice as its subject matter. Which determination of choice can be viewed as good is stated by the moral law alone. On the one hand, this law subjects all our volition to a negative condition. A morally permitted end is only one which can be realized in an action whose maxim is desirable as a general law. On the

other hand, the moral law determines certain material ends which the agent should rate higher than his private ends. Accordingly, the agent should have maxims that correspond to these material ends. According to Esser, the moral law thus governs a process of deliberation determining which maxim in a particular situation is morally justifiable and which end can be thought to be valuable. It is not the case that we find ourselves confronted with conflicting abstract values. Kant's ethics provides us with a criterion that releases us from the absolute claim of abstract values by indicating which maxim is permitted under the given circumstances so that it is of moral value. It is thus the maxim which can be morally judged. Esser suggests that a maxim is best understood as hypotheses of the basic moral attitude of one's own acting and that of others. An agent can determine his will and his maxim in any situation. What is not at his disposal is the concrete power to realize certain ends under all circumstances. It is not the individual action that is crucial for assessing the morality of a person, but solely which basic principles are revealed by the actions of that person.

Esser shows that even if there is a rational resolution of moral conflicts, there remains room for emotional reactions such as regret. What is regretted, however, is not that a moral principle has been breached. Rather, there are other reasons to render such regret rational, such as the loss of something valuable, terrible circumstances, tragic consequences, or the agent's own helplessness.

According to Esser, Kant claims neither that morality is capable of disposing of tragic conflicts, nor that it can save us from the consequences of hard choices. What morality can do is to inform the agent in each case what may be morally assigned at all. But it is up to the agent's judgment to find out what is at his disposal for realizing his ethical maxim.

The thread that persistently runs through all the papers in this collection is that Kant's moral theory differs from ancient and contemporary virtue ethics. But his later ethical writings reveal that themes of virtue ethics play a more prominent role than typically thought. Many contributors, however, also share the view that Kant remains faithful to the priority of the right. What we are morally required to do remains grounded in the moral law. But this does not preclude the fact that moral excellence, the motivations of moral agents, and their ends are important as well.

The resources Kant provides in his later ethical writings help to show that many of the concerns raised about his moral theory can be assuaged. Despite his view that it is the moral law that provides the sole principle

for our behaviour, his notion of imperfect duties helps to show that his ethics is neither morally too demanding, nor too impartialist and overly rigoristic.

The essays here suggest that Kant's ethics, to be sure, is not to be assimilated into virtue ethics. It is not a conception of the ideal person that provides the foundation for morality. But Kant's later writings help us to see that virtue is a core element in his ethics, precisely because it helps us to do our duty. It is the aim of this collection to provide further insight into the role virtue plays.

There is more to be discussed, to be sure, but one lesson to be drawn from these contributions so far is that it is not a contradiction in terms to speak of Kant's ethics of virtue, even though it is not to be confused with virtue ethics.[3]

3 I am grateful to Christian Budnik, Ellen Coughlin and Jörg Schroth for extensive written comments. Many thanks also to Christine Bratu and Tobias Wille for editorial assistance.

Kantian Virtue and 'Virtue Ethics'

Thomas E. Hill, Jr.

The reputations of various systematic ethical theories rise and fall over time, for good and bad reasons. Some rise like splendid towers only to be demolished by withering attacks from critics. Others like castles on hilltops were apparently too formidable to be destroyed by frontal assault, but their defenders abandoned them in search of newer, less cumbersome quarters, leaving only their impressive ruins behind as tourist attractions.[1] We can hope that, as in the history of medicine, there is progress in these changes guided by improved understanding and wise reflection; but we may fear that, as with fads and fashions, the changes are driven by enthusiasm for whatever is novel. Despite our hopes, progress is not guaranteed.

Similar remarks might be made of major *approaches* to ethics as well as particular ethical theories. For example, theological approaches were once dominant but lost ground to various forms of intuitionism and utilitarianism. Attempts to restrict ethics to the "metaethical" analysis of moral language have largely been abandoned in favor of more holistic approaches, as illustrated, for example, by the influence of Rawls' theory of justice. Today, it seems widely agreed, normative ethics is largely dominated by several competing approaches, including prominently utilitarianism, Kantianism, and virtue ethics. Utilitarianism has undergone many subtle transformations in response to criticisms and still survives with many adherents especially in Great Britain. Broadly Kantian approaches to ethics have been developed in a variety of ways that have earned new respect and attention for their source; but simultaneously a diverse movement commonly called "virtue ethics" has gained popularity especially in its opposition to other approaches. In my discussion today I shall be concerned with the last two of these trends, but primarily the first. More specifically, I shall discuss first Kant's conception of moral virtue, its place in his overall ethical theory, and some questions that it raises. Then, all too briefly, I shall consider how Kant's ethics, or at least

1 The imagery here comes from Isaiah Berlin's lectures at Oxford in the early 1960's.

an expanded and supplemented Kantian ethics, might respond to certain criticisms and alternatives often associated with "virtue ethics".

My aim is to explore a wide range of related topics, to offer suggestions regarding the interpretation and development of Kant's ideas, and to raise questions about how these compare and contrast with some ideas in another tradition.[2] With apologies to scholars of both traditions, here I can only sketch various positions with broad strokes, passing over various complexities and controversies that may be important in another context. My suggestions often presuppose interpretations and arguments that I discuss more fully elsewhere, but even so they remain tentative as well as incomplete. Here my hope is just to call attention to some features of Kant's ethics that emerge mostly in his later writings, to address some problems, to warn against some misunderstandings, and to facilitate further discussion of the relative merit of different approaches.

I. The Apparent Conflict: Basic Points and Exaggerations

It may be useful, especially for those not already familiar with the controversy, to begin with a loose description of Kant's ethics and virtue ethics as they are often conceived – or caricatured. In an extreme (distorted) picture, Kant's ethics is seen as exclusively focused on moral duty, the sole issue being how I must act and not what sort of person I should be. All duties are supposed to be derived from reason alone

2 Earlier versions of this essay were discussed at the University of Tennessee, the University of Cincinnati, the University of California, Riverside, and the University of California, Los Angeles. A lecture with similar content to a few sections of this present essay was given at The Catholic University of America. (That lecture, entitled "Kant on Weakness of Will," will be published by The Catholic University of America Press in a series of essays on weakness of will, edited by Tobias Kaufmann.) I am grateful to participants in those discussions for their helpful comments. I cannot thank individually all the Kant scholars from whose work I have learned, but among the best accounts of these matters, in my view, is Robert Johnson's "Weakness Incorporated", 1998, in: *History of Philosophy Quarterly*, vol. 15, no. 3, pp. 349–367. My reflections agree with his excellent discussion on most important matters, though I do not rely in the same way on a distinction between "motivation" and "values". Also at times Johnson seems to treat weakness of will as a kind of inability. For example, "The weak person we are interested in here is not simply a person who *drops* and *adopts* principles too easily, but is rather one who *cannot* live up to those she adopts and does not drop" (p. 360, my italics in the second clause).

independently of any assumptions about human feelings or other empirical data about the human condition. Ethical standards are not scalar ideals of excellence, which we may "live up to" or "fall short of" to various degrees. Instead, moral principles are always definite and non-scalar, or all-or-nothing: either you obey the strict moral commands or you don't. Morality is completely rule-governed, like an all-encompassing legal code consisting of rigid, exceptionless laws that specify for all cases what is forbidden and what is required. The only good motive is devotion to duty: to do duty for duty's sake. Other motives are morally on a par: being moved by compassion, for example, is *no better* than being moved by malice. We should only do our duty from respect for moral law, not from habit, not to emulate virtuous persons, or even not to do what is noble and honorable. A moral life must be entirely governed by strict, inflexible, moral rules that bind everyone, regardless of circumstance, without exception. Moral rules completely ignore special relationships between individuals and differences in historical and cultural contexts. Knowing what is right to do is obvious to even the most ignorant, unreflective, and corrupt persons: we never need to weigh expected consequences and virtue is not needed to discern the right thing to do. Acting rightly and having good traits of character are equally within the power of every minimally competent adult: *luck* has nothing to do with it. Justice requires judicial punishment and judgments of blame according to inflexible, exceptionless rules that serve the basic aim to make each person suffer according to their inner moral deserts.

By contrast, "virtue ethics" is sometimes conceived as follows. Its focus is *character* or what one should *be*, not how one should *act.* Virtues of character are derived entirely from the *study of human nature*, especially what enables us to flourish as human beings. This is determined entirely by our *natural end* (or *telos*) and our natural powers and vulnerabilities. All moral standards are scalar ideals of excellence, to which we may approximate more or less. Rules and principles are of little use outside of a legal system. They have no authority independent of judgments on particular cases, and focusing on them inevitably leads to the neglect of morally relevant features of the case at hand. There are many virtuous motives, and the appropriate kind and degree depend on the context. Concern for the noble and honorable, for example, is a general mark of a virtuous person. Compassion and many other specific motives are good but only when their degree, time, place, and manner of moving us are fitting to the particular circumstance, as judged by a practically wise person. Acting from good habits should be encouraged; and to develop

such habits, we should emulate those who are known paragons of virtue. Friendships and family relations are outside the scope of duty and moral rules. The standard of right action is to do whatever a thoroughly virtuous person would do – so the *right* is completely determined by prior judgments about what is *good* in particular cases. Virtue and knowledge make a person better able to judge what is right and good, and both depend heavily on how *lucky* one is in one's natural endowments, social circumstances, and friends and mentors. A good life requires sensitivity to the rich variety of personal relationships and historical contexts. Moralistic blame and punishment belong to an untenable legalistic model of ethics that values uniformity, abstract rules, and vengeance more than human flourishing and good character.

My sketch of Kant's ethics is a caricature combining mistakes and distortions that probably no one believes altogether. It reflects significant themes in Kant's writings but exaggerates and misrepresents them in unfortunate ways and omits important qualifications. It also fails to distinguish what is deep and basic in Kant's moral theory from ideas that are less so. My sketch of *virtue ethics* is perhaps less a caricature than an agglomeration of points from different versions of virtue ethics, or points offered from time to time by those inclined to favor a "virtue ethics" approach. A thorough treatment would need to clarify the apparent differences among these views, to sort out the deep contrasts from exaggerations and misunderstandings, and to identify precisely which moral theorists, if any, are committed to the views.[3] This essay attempts to take a few preliminary steps in this direction, at least concerning Kant's ethics.

In contrast to most philosophers before and after him, Kant treats *virtue* as a kind of strength of the will to do what is right. Virtue is more than having good intentions, and we need to develop it over time. We have a duty to try to develop virtue, but we are also responsible if, lacking

3 Serious discussions of virtue representing contemporary versions of a broadly "virtue ethics" approach, as distinct from more popular and implausible representations, can be found in the following: Philippa Foot, 2003, *Natural Goodness*, Oxford: Oxford University Press; Stephen Darwall (ed.), 2002, *Virtue Ethics*, Oxford: Blackwell Publishers; Roger Crisp/Michael Slote (eds.), 1997, *Virtue Ethics*, Oxford and New York: Oxford University Press; Rosalind Hursthouse, 1999, *On Virtue Ethics*, Oxford: Oxford University Press; Michael Slote, 1992, *From Morality to Virtue*, Oxford: Oxford University Press; and Christine Swanton, 2005, *Virtue Ethics: A Pluralistic Approach*, Oxford: Oxford University Press.

sufficient virtue, we do wrong through weakness of will. We have particular virtues insofar as our will to fulfill various specific duties is strong, but it is not the implications of Kant's position for particular virtues (and vices) of character that is most interesting. Apart from that, Kant's ideas raise questions about moral responsibility and strength and weakness of will that present special problems of understanding, even though (arguably) in some respects Kant's ideas reflect familiar common sense views.

My discussion will be divided into three main parts. Part II reviews briefly basic ideas in Kant's *Groundwork for the Metaphysics of Morals* about methodology, a good will, morally worthy acts, categorical imperatives, and the supreme moral principle. Part III concerns ideas that Kant develops in his later ethical writing, notably two senses of "will", virtue as more than a good will, and a distinction between weakness, impurity, and depravity. Central questions here are: How can we understand moral weakness of will as a deficiency of virtue, not a disability, as culpable but better than impurity and depravity? And how does Kant's idea that we act on *maxims* fit with his metaphors of strength and weakness of will? In Part IV, I consider implications regarding four topics on which Kant's ethical theory might be contrasted with certain views associated with "virtue ethics": 1. the need for moral rules, 2. sensitivity to particular contexts, 3. morally good motivation, and 4. the standard for right action.

II. Some Basic Themes of the Groundwork

Let us review briefly ideas that Kant introduces in his *Groundwork for the Metaphysics of Morals.* These set the stage for his discussion of virtue in his later work, *The Metaphysics of Morals.*

II.1. Methodology

To begin, the method appropriate to *foundational* questions of ethical theory, Kant thought, was an a priori method that did not rely on empirical studies of human psychology. There were two main reasons: first, the way to "discover" and confirm a principle as the supreme moral principle, he thought, was to *analyze the common idea of duty*, and this is a matter of deep conceptual analysis, not empirical investigation. Second,

foundational ethics needs to resolve the question *whether commitment to morality is rational*, and empirical studies cannot determine what constitutes practical rationality.[4] Although many philosophers today doubt Kant's conclusions, many still accept his main methodological point: that basic questions about the presuppositions of our concepts and the nature of rationality call for philosophical reflection and cannot be resolved by empirical studies.

Unfortunately, this sensible point has been confused with another, utterly untenable idea: namely, that we can judge what is right to do in a particular situation without relying on any empirical facts at all. Although Kant's incautious remarks sometimes contribute to the confusion, the idea that empirical evidence is irrelevant to moral judgments about particular cases and specific (mid-level) principles is both absurd and contrary to Kant's own moral arguments. We must rely on empirical facts even to identify an act as a "lying promise", "a suicide", "an act of revolution," or an act of another kind. Moreover, as Kant acknowledged, his own substantive principles in *The Metaphysics of Morals* about political authority, punishment, international relations, self-perfection, beneficence, gratitude, respect, and friendship presuppose general assumptions about the human condition. He argued that some acts are morally wrong, whatever the consequences, but he knew, as any sensible person does, that at some level *consequences matter.*[5] He held, for example, that we have

4 These two tasks of fundamental moral philosophy in the *Groundwork* were to "seek out" and "establish" the supreme principle of morality. Immanuel Kant, *Groundwork for the Metaphysics of Morals*, ed. by Arnulf Zweig and Thomas E. Hill, Jr., 2002, Oxford: Oxford University Press (hereafter abbreviated as *G*) 193 [4: 390].

5 Consequentialists hold that *nothing but* consequences, and perhaps a few "side constraints", ultimately matter, and Kant is no consequentialist in this sense. Kantian moral reasoning to particular moral conclusions, however, is complex and consequences are potentially relevant at certain stages though constrained by certain principles and procedures. It should be noted that any substantive moral principle that constrains choice, even if justified by a procedure that takes consequences into account, will exclude in the particular case consideration of facts (including consequences) about anything but the criteria for judging the case as falling under the principle. As an absolute principle "Never lie," whatever its grounds, prohibits considering any facts beyond the criteria for determining what would be a lie in the case (among which, for example, is that my saying certain words now will likely cause the other person to believe what I think is false). Even a qualified principle, such as "Be truthful, except when deception is necessary to save lives and is directed to a person making an immoral threat (and other specified conditions)" says, in effect, "In these circumstances, you must not

duties to preserve our health, promote the happiness of others, and work towards world peace. One cannot even begin to do these things without estimating the probable consequences of our acts.

II.2. A good will: the core feature of a morally good person

After describing his methodology, Kant begins his actual investigation with the famous assertion that *only a good will is good without qualification.* This thesis, he implies, is inherent in our ordinary rational cognition of morality. The basic point here is that what a rational and good person would seek and preserve, no matter what else must be sacrificed, is a *good will* – a basic commitment to do what is right, whatever the circumstances. Without this no one can be a morally good person. Other traits can be useful and admirable but are not essential or sufficient for basic moral goodness. Kant emphasizes that a good will expresses itself only in acts from duty, but insofar as a good will is a stable disposition over time one can *have* a good will even when acting from other motives. The moral goodness of a person, then, is not determined by *how many* "morally worthy" *acts* the person does from duty. That may depend in part on the moral challenges that the person faces. A person blessed with few temptations is not necessarily a worse person than others because she faced fewer occasions that called for explicit thoughts of duty.[6] A good will can serve as a *back-up* motivational commitment even when we are doing things for other reasons. We play games for fun, but a good person is ready to stop to help others in emergency situations.

tell a lie regardless of *further* consequences." That substantive moral principles exclude (further) consequences in *this* way is not a radical Kantian thought but a thought that few, if any, who acknowledge genuine moral principles (in contrast with act-consequentialist "rules of thumb") would deny.

6 This interpretation may be controversial because Kant seems to imply (and *at times* it seems true) that to do what is right from non-moral motives is a moral failing. Consider a teacher who never gives her favorite student (who happens to deserve high marks) a low grade *simply because she likes him*, with no thought of her duty as a teacher. At least in the case of various negative duties, however, there is no adequate reason to suppose that Kant was committed to the implausible idea that on each occasion on which one satisfies the requirement from a non-moral motive one is morally deficient. For example, well-contented people who pass up many unwanted opportunities to commit suicide because they want to live need not be considered guilty of anything so long as, with a good will, they stand ready to avoid suicide when tempted and when it would be an escape from responsibilities.

Often we help from love, not duty, but a person with a good will is ready to help, if necessary, even when wishing to do something else. Kant argues that the freedom presupposed in moral agency entails that whether or not we have a good will is up to us, not empirically determined. By contrast, our desires, our feelings, and the *effects* of our actions are often beyond our control.

II.3. Morally worthy acts and acts commendable in other respects

Kant turns next from evaluation of persons to evaluation of acts. His central point is well-known and often criticized: *assuming that an act accords with duty, it has moral worth if and only if it is done from duty.*[7] To act from duty is to do something because it is right, not because it is profitable or immediately agreeable. The more difficult question is what is "moral worth"? The key is its connection with the essential feature of a good person: *morally worthy acts manifest a good will in action* whereas other acts do not. This is the feature that Kant was concerned about for purposes of his argument: his aim in *Groundwork* I was to "seek out" the supreme moral principle by asking, 'On what principle is a good person acting when moved by the attitude or commitment essential to morally good persons?'[8] Kant's initial answer is: a person with a good will, when acting as such, does what is right, regardless of rewards and punishments. This account is later expanded: for human beings, acts motivated by a good will reflect the agents' *respect for the moral law* rather than their desire for a particular end. Finally the principle of a good will in action is said to be "*Conform to universal law*", which is all too quickly transformed into the famous first formulation of the Categorical Imperative: *"Act only on that maxim by which you can at the same time will that it should become a universal law."* The details of the argument are not necessary here. The point is that Kant's famous claim about morally worthy acts is not presented as a complete theory about how to assess the moral value of acts, still less a complete theory about how to assess character. To the

7 *G* 199–201 [4: 397–9].

8 At this point Kant's strategy of argument bears some resemblance to the idea in virtue ethics that we need to determine what is right to do by considering what a good or virtuous person would do. A significant difference is that Kant seeks a general *principle*, not particular judgments, by this strategy. Also his conception of a person with a good will is apparently different from the "thicker" ideas of a virtuous person typical of virtue ethics.

contrary, it is a step, arguably even an unnecessary step, in an early argument to identify a fundamental criterion of right action.

The practical, forward-looking message of Kant's claim is that one should strive to be the sort of person who takes the fact that something is morally required as a sufficient reason for doing it. This does not mean that we should be always thinking of duty, especially not that we must always be focused on Kant's abstract principles. Kant suggests that traits of character other than a good will, such as moderation and self-control can be (morally) commendable provided the agent also has a good will.[9] Later he adds that it is "an indirect duty to cultivate the compassionate natural (aesthetic) feelings in us," and so such compassion has moral value at least as a trait that good persons must seek to cultivate.[10] Even if they are not done from duty, acts that manifest these traits may also be commendable as such, provided the agent has a good will (as a standing commitment and "back-up" motive). Kant does assert, however, that *only acts from duty* have the special kind of value he calls *moral worth*. To reconcile these ideas, we must distinguish between two perspectives from which we may evaluate the moral value of acts. One perspective focuses on the agents' commitment (or will) to act as they should; the other focuses on the agents' success in cultivating and exemplifying character traits that help us to live a moral life and achieve our moral ends. Because they manifest a good will, *acts from duty* are worthy of *esteem as instances of the agents' willingness* to live up to moral standards, even when otherwise inclined. Because they stem from character traits that, for moral reasons, everyone should *try* to cultivate, *acts that demonstrate moderation, self-control, and even sympathy* may be *commendable as exemplifying ideals that we should strive for*, even if success in cultivating these traits is not entirely up to us. Once we make the distinction, we see that, though Kant insisted that a good will and acts from duty have a special kind of moral value, he agreed with the common opinion that other character traits and acts can also have a kind of moral value that is undeniably important.

9 *G* 195–6 [4: 393–4].

10 Immanuel Kant, *The Metaphysics of Morals*, trans. and ed. by Mary Gregor, 1996, Cambridge: Cambridge University Press (hereafter abbreviated as *MM*) 205 [6: 457].

II.4. Moral and non-moral reasons

Kant begins his argument in *Groundwork* II by considering the nature of *"ought" judgments in general* and then what is distinctive about the *idea of moral duty.* This chapter, like the previous one, aims to "seek out" the supreme principle of morality by an analytic mode of argument. It aims to discover the deep sources of morality presupposed in the common idea that we are subject to moral requirements that are distinct from the requirements of efficiency and prudence. The idea of duty serves as the second of two *starting points in Kant's initial investigation*, but it is the idea of *autonomy, not duty,* that he finds to be *the source of morality.* A good will and duty are *first for expository purposes*, but rational autonomy is *primary for purposes of deep explanation and justification.*

"Ought" judgments in general, on Kant's analysis, purport to express "objective principles" that are "necessitating" for "imperfect wills."[11] In other words, they purport to say what we have good reason to do, and so "must" do, even though we might not. If justified, a judgment that one "ought" or "must" do something affirms an *imperative* – that is, a requirement of *reason* expressed in a form appropriate for those who do not automatically do what is rational.

Some "ought" judgments are non-moral, for example, "You ought to practice shooting free throws" and "You ought to stop and smell the roses." Here the reasons for the prescriptions are dependent on the prescribed act serving as a means to an end the agent happens to want: for example, to improve as a basketball player and to be happy. Other "ought" judgments are moral: for example, "You ought to keep your solemn promises" and "You ought not kill people for profit." Here, Kant argues, the predominant reason is not that the prescribed act or omission serves the ends you happen to have. And you cannot render the requirement inapplicable to you simply by altering your ends. In Kant's terms, the first "ought" judgments are *hypothetical imperatives*; the second (moral) "ought" judgments are *categorical imperatives.*[12] If they are genuinely rational requirements, they derive ultimately from a fundamental principle of rationality. In the first case, the principle seems to be "One ought to take the necessary means within one's power to the ends one wills, or else revise or abandon the ends." Commentators sometimes call this "the Hypothetical Imperative." In the second (moral) case, the

11 *G* 214–5 [4: 413–14].
12 *G* 215–18 [4: 413–417].

fundamental principle is what Kant calls "the Categorical Imperative", and he argues that it can be expressed in several ways.

Three formulations should be mentioned, though they are not my main focus here. Famously, the first says: "Act only on that maxim by which you can at the same time will that it should become a universal law."[13] In other words, it is wrong to act on basic policies that you cannot reasonably choose as policies that others may (or should) follow. The second, equally influential, formulation says: "Act in such a way that you treat humanity, whether in your own person or in any other person, always at the same time as an end, never merely as a means."[14] On one reading, this says essentially that we should treat all persons in ways consistent with their dignity as rational beings. Dignity is "an unconditional and incomparable worth," above all price, and without equivalent. A third formulation is "Act on the maxims of a universally law-giving member of a merely possible kingdom of ends."[15] All too briefly, the point is that the general moral requirements on everyone are those that rational and autonomous persons would prescribe or "legislate" for all from an appropriate legislative perspective that is modeled on an analogy with an ideal commonwealth where everyone is both author and subject of the laws.

These topics – categorical and hypothetical imperatives and the formulations of the fundamental moral principle – are the most often discussed features of Kant's ethics, but for comparison with virtue ethics I want to highlight just two points, often neglected. *First*, although Kant notoriously endorsed certain strict prohibitions (for example, against lying) as unqualified, inflexible requirements, arguably the strictness of his personal convictions on these particular matters is neither inherent in the idea of a categorical imperative nor derivable from his fundamental moral principle. By definition, a categorical imperative is a principle of the form "Ought never…" or "Always ought…" and we cannot avoid its demands simply by revising our ends. But the concept of a categorical imperative allows that the principle itself may contain built in qualifications, and Kant implicitly acknowledges this. For example, Kant apparently thought that it is a categorical imperative never to kill another human being except in self-defense, a just war, or legally authorized execution. Moreover, some categorical imperatives prescribe

13 *G* 222 [4: 421].
14 *G* 230 [4: 429].
15 *G* 239 [4: 439]. See also *G* 233–35 [4: 433–4].

the adoption of general ends, such as the development of one's talents and the happiness of others. These requirements do not say definitely when, how much, or in what ways one should pursue the moral end, but rather leave details to judgment and choice in particular contexts.

Second, as Kant analyzes moral imperatives, they combine two distinct elements: (a) an affirmation that there are sufficient, overriding reasons to act in a certain way and (b) an implication that we are constrained to do the act, contrary to what we may want. The former is inherent in the idea that categorical imperatives express "objective principles." The latter is implicit in the rather demanding vocabulary: "must," "ought," "imperative," and "necessitating." Now, for various reasons, Kant concentrates on cases where both elements are apt: there is a rational requirement and we are disinclined to fulfill it. But the elements can come apart in other cases. For example, those who simply acquire their moral beliefs from authorities may accept that *they are constrained* to do or avoid certain acts though *they have no awareness of the reasons* for doing or avoiding them. Alternatively, we may be well *aware of the reasons* but *lack any sense of constraint* – because we are not in the slightest tempted to do otherwise on the occasion. These distinctions are important because they open up the possibility that we might often act from our recognition of the good moral reasons behind a categorical prescription without the sort of reluctant, unhappy, self-sacrificing attitude commonly associated with the idea of "acting out of duty." Arguably, even by Kant's criteria, such acts should be counted as "morally worthy" because they manifest a good will insofar as they are motivated by good moral reasons. Kant himself did not address the possibility of such unconstrained moral acts, perhaps because he was so pessimistic about human inclinations and so keen to emphasize that a moral life can be hard and demanding. But the possibility is not ruled out by the basic elements of his theory.

III. Kantian Virtue

A good will is the essential feature of a morally good person, but virtue is more than a good will. In *The Metaphysics of Morals* Kant gives an account of virtue as strength of moral will, something that requires time and practice to develop.[16] A good will is a fundamental commitment to

16 *MM* 145–6 [6: 380], 148–9 [6: 383–4], 151 [6: 387], 153 [6: 390], 156–7

doing what is right, despite obstacles, but a good will may be weak. A virtuous person has a will that is both good and strong. It is not only able but fully ready to overcome inclinations to act contrary to duty, and it does so effectively and without wavering before temptation. One has particular virtues insofar as one has a strong will to fulfill the various moral requirements of justice, truthfulness, self-improvement, beneficence, gratitude, respect, and so on. The common core of virtue, however, is a good and strong will to do what is right.

The basic idea here may seem familiar, but to understand it in the context of Kant's philosophy requires some background. Once the idea is placed in this larger context we can see what is special about it. In particular, Kant's idea of strength of will contrasts sharply with empiricist conceptions, and it raises special problems.

The background includes prominently Kant's ideas of inclination, practical reason, and will (in two senses). Briefly, we have *inclinations* by nature and social influence. These are feelings and dispositions of which we are aware empirically – impulses, desires, aversions, and steady interests.[17] They are not all self-directed even though Kant classifies maxims based on them under "self-love."[18] We can try to cultivate or rid ourselves of particular inclinations, but our ability to do so is limited. Inclinations and feelings are often obstacles to clear thinking and right acting, but in themselves they are neither good nor bad.[19] They provide impetus to action, but they are not to be seen as sufficiently determining causes.

As human beings we also have *reason*, theoretical and practical.[20] *Theoretical reason* enables us to discern facts about what there is (the world as it appears) through logic, fundamental concepts of understanding, science, and everyday learning processes. *Practical reason* enables us to discern the fundamental principles of what we ought to do. It can guide deliberation, judgment, and choice, not merely telling us what is

[6: 394–5], 163–8 [6: 404–10], 223 [6: 479–80]; Immanuel Kant, *Critique of Practical Reason*, trans. and ed. by Mary Gregor, 1997, Cambridge: Cambridge University Press (hereafter abbreviated as *CPrR*) 72 [5: 84–5] and 107 [5: 128].

17 See, for example, *G* 215n [4: 414n].

18 *CPrR* 19–20 [5: 22–3].

19 Immanuel Kant, *Religion within the Boundaries of Mere Reason*, ed. by Allen Wood and George di Giovanni, 1998, Cambridge: Cambridge University Press (hereafter abbreviated *R*) p. 46–7 [6: 20–2] and 57 f [6: 34–5]. But see also *G* 229 [4: 428].

20 *CPrR* 2–5 [5: 4–6].

right but disposing us to act accordingly.[21] In his later work Kant distinguishes two senses of *will* – a power of choice (*Willkür*) and a rational legislative will (*Wille*).[22] Neither is an empirically observable "thing," but we need ideas of both to make sense of moral deliberation and judgment. Rational legislative will (*Wille*) is an aspect of practical reason, metaphorically represented as the will of an authority (reason), which is necessarily a part of each rational agent, and this authority "legislates" to the agent universal principles of conduct. As rational human agents we cannot help but feel respect for this authority, and yet its prescriptions often conflict with our inclinations. The result is not a mere battle of opposing forces, for we also have a power of (free) choice (*Willkür*). When inclinations and rational dispositions conflict, we must presume that we are able to choose and are responsible for our choices.[23]

Our actions are interpreted as based on *maxims* reflecting our intentions, ends, and (sometimes) underlying reasons.[24] At least for morally significant cases, we must view ourselves as acting on freely chosen rationales, not merely exhibiting behavior that is the product of causal forces. When we act from inclination, we are not causally determined by forces within us but instead we "incorporate" the inclination into a maxim on which we freely choose to act.[25] Or so, for moral purposes, we must suppose. For example, we must see ourselves as freely adopting the policy or intention to satisfy our lust rather than supposing that the lust causally determines our behavior. What is presupposed here is *freedom in a negative sense* – an ability to act without being causally determined or unavoidably motivated by empirical factors beyond our control ("alien causes").[26] This, Kant argues, is inseparable from *autonomy of the will*, or freedom in a positive sense.[27] The core of this idea is that as rational agents we are able and disposed to guide our conduct by rational principles of conduct that are not desire-based, do not merely prescribe taking means to our ends, and yet are products of our own reason. As human beings that have both inclinations that often oppose reason and a (negatively free) power of choice, unfortunately we

21 *CPrR* 29 [5: 32].
22 *MM* 13 [6: 213] and 18 [6: 226]. Gregor translates *Willkür* as *choice* and *Wille* as *will.*
23 *G* 247–8 [4: 447–8]; *CPrR* 37–44 [5: 42–50]; *MM* 18–9 [6: 226–7].
24 *G* 202n [4: 402] and 222n [4: 421]; *MM* 17–18 [6: 225–6].
25 *R* 48–9 [6: 23–4].
26 *G* 246 [4: 446].
27 *G* 246 [4: 446], 231–5 [4: 431–4] and 240–1 [4: 440–1].

do not always express our autonomy by following those rational principles. Nevertheless only mistaken moral theories deny that moral agents have it.[28]

Particular maxims or rationales may be good, bad, or indifferent, and Kant acknowledges maxims of different levels of generality. The *Groundwork* provides examples of relatively specific maxims regarding suicide, lying promises, neglect of talents, and helping others.[29] Later Kant refers to fundamental, life-governing maxims reflecting one's deep and persistent priorities when morality and self-interest conflict.[30] Here our choices are more limited: the only good choice is that of a good will, which has a basic commitment to prioritize morality over self-interest. Any other choice is evil. This is the background for Kant's distinction between human tendencies to weakness, impurity, and depravity.[31] Human beings, in Kant's view, are never fiendishly evil in the sense of willing to do evil for evil's sake.[32] Moral *depravity* is the systematic subordination of morality to self-love in one's fundamental, life-governing maxim.

Impurity consists in having an unstable, ultimately incoherent mixture of conflicting elements in one's basic maxim, manifesting a failure to make an unqualified commitment to either morality or self-love. A person with an impure moral will may, for example, adopt as a basic maxim "I will do what is morally required, unless the cost is too high."[33] Neither a morally *depraved* person nor a morally *impure* person is

28 *G* 241–45 [4: 441–4].

29 *G* 222–4 [4: 421–3].

30 *R* 45–60 [6: 19–39]. See also *CPrR* 17–24 [5: 19–26].

31 *R* 52–4 [6: 29–31]. Kant introduces the terms to describe "propensities" to evil in human nature in contrast to several "predispositions" to what is basically good, *R* 50–2 [6: 26–28]. I extend the terms to characterize those who have voluntarily incorporated these propensities into their basic, life-governing maxims.

32 *R* 58 [6: 35]; *MM* 160 [6: 399–400] and 208 [6: 461].

33 The incoherence of this position stems from the fact that Kant understands a "moral requirement" (or duty), when fully stated, as an unconditional imperative, expressible as "Do (or avoid) X, regardless of how it affects your contingent interests." Thus the mixed maxim would be, in effect, "I will do what is morally required, *regardless of how it affects my interests* but *only if it does not too much damage my interests.*" Note, however, that this point is compatible with thinking that sometimes effects on one's own contingent interests are relevant in deciding the specific content of a moral requirement, for example, whether one must give absolutely all of one's goods to the poor.

committed to the unconditional moral law as a life-governing maxim, and therefore neither has a *good will* in the primary sense of the *Groundwork*.[34]

A morally *weak* person, by contrast, has a fundamentally good will but lacks virtue. That is, such persons incorporate the moral law into their life-governing maxim, giving it in principle priority over self-love, but they deviate from the moral law on particular occasions because they have not developed a sufficiently *strong* will to do what is right. As Kant says, the weakness or "frailty (*fragilitas*) of human nature is expressed even in the complaint of the Apostle: 'What I would, that I do not!' i.e. 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 (*in thesi*), is subjectively (*in hypothesi*) the weaker (in comparison with inclination) whenever the maxim is to be followed."[35]

It is moral weakness that is especially important for understanding how *virtue* is more than a *good will.* Those who in their basic life-governing maxims yield to the propensities to impurity or depravity lack a good will: they have no commitment to moral standards as unconditional requirements. As moral agents, both inevitably have a rational legislative will (*Wille*) and so feel respect for morality, and therefore they are liable to suffer for their wrongdoing the self-disapproval that Kant pictures metaphorically as a sentence in the inner court of conscience.[36] Unlike the impure and the depraved, morally weak persons have a good

34 In the *Groundwork* Kant does not explicitly distinguish a *good will* and *virtue*, but it is clear that a "will" (or resolve or firm intention) to act as moral law requires is at least the core of both. At *G* 196 [4: 394] Kant describes a good will as not "merely a wish but a summoning of every means in our power." This implies that that a person with a good will exerts (or at least intends to to exert) every effort to follow moral requirements. This is characteristic of virtue, but it is unclear at this point whether it necessarily includes the virtuous person's *strength or fortitude of will*, a developed moral resolve that is persistent and effective over time. At *G* 253–4 [4: 454–5] Kant writes of a "malicious villain" as "conscious of possessing a good will which, on his own admission, constitutes the law for his evil will as a member of the world of sense – a law of whose authority he is conscious even while transgressing it." At this point "a good will" seems to refer to what Kant later identifies as a rational legislative will (*Wille*) rather than an adoption through a free power of choice (*Willkür*) of a basic, life-governing maxim to follow the moral law. A malicious villain might have a "good will" in that sense (*Wille*) even though morally depraved (as opposed to morally weak or virtuous).

35 *R* 53 [6: 29].

36 *MM* 188–91 [6: 437–40].

will – the basic life-governing maxim that expresses their most fundamental choice regarding morality and self-interest places morality above self-interest, unconditionally, in case of conflict. They have a good will, then, but lack virtue because their will to live by this basic commitment is weak. Virtue requires a good and strong will. But how are we to understand this?

Because we are responsible for our basic, life-governing maxims, these must be represented as due to our free power of choice (*Willkür*). Those who are *depraved* have elected (by *Willkür*) to subordinate the prescriptions of their rational legislative will (*Wille*) to self-love and those who are *impure* have chosen to accept only *qualified* conditional versions of those prescriptions. Thus neither has a good will. The *morally weak* have a basic commitment (through *Willkür*) to follow the full, unqualified prescriptions of their rational legislative will (*Wille*), but they fail to live up to this commitment on some particular occasions. (If they failed readily and regularly, however, we would naturally doubt that they had the basic commitment.) They have a good will, but, as we say, it is weak. By contrast, the fully *virtuous* have the same basic commitment (through *Willkür*) but, unlike the weak, they do not deviate from it on particular occasions. This is because they have a developed strength of moral will to overcome the obstacles of inclination. Their will is both good and strong.

The question remains: how are we to understand strength and weakness here? In Kant's view, those who do wrong through moral weakness, though better than the impure and the depraved, are nevertheless directly responsible for their wrongdoing.[37] Their weakness, then, is not a disability, like underdeveloped or atrophied muscles. Thus the special strength of moral will that virtuous persons have (and the morally weak lack) cannot then be literally the capacity or ability to do right despite contrary inclinations because, having a free power of choice, even the morally weak have that capacity. We cannot free ourselves from responsibility for serious wrongdoing by citing the weakness of our will to do right, as if this is a disability such as a physical weakness that causes us

37 This is entailed by Kant's repeated insistence that we can do what we ought. Even the innate human "propensity" to moral weakness, Kant held, must be seen as chosen, despite any apparent empirical evidence to the contrary. *R* 52–55 [6: 29–32]. Insofar as moral weakness is merely "want of virtue," as evidenced by neglect (rather than principled rejection) of imperfect duties (beneficence, etc.), it is not "vice" but "lack of moral strength" (of "resolution") and assessed as "a deficiency in moral worth." *MM* 153 [6: 390].

to fail in an attempt to rescue someone drowning. Kant suggests that we should take a charitable view towards others' weakness, but, strictly speaking, his theory implies that we must suppose that they too have the capacity and freedom to do whatever morality requires.

Kant also rejects the position (for example, of Thomas Hobbes and David Hume) that choices between good and bad are ultimately to be explained by whether the agents' broader-based desires and sentiments were causally more efficacious than their more immediate, impulsive ones. Moral weakness, then, is not simply a matter of having dispositions to act rightly that have *greater force* than our dispositions to act wrongly. By his own admission, indeed insistence, the exercise of agency (or "will") cannot be explained or comprehended theoretically, either in empirical or metaphysical terms.[38] So strength and weakness of will, as conceived in moral discourse, must be treated as metaphorical (or "noumenal") ideas needed and valid only for practical purposes. Thus we cannot expect references to strength and weakness of moral will to provide a "nuts and bolts" type of causal explanation. These are not terms that fit into any empirical science, though Kant grants that any "phenomena" associated with morally significant acts, indeed any acts, can also be viewed from an empirical perspective that presupposes causal determinants.

Here is one way we might interpret or reconstruct Kant's position. The morally weak, unlike the virtuous, act contrary to their own basic commitment (good will) to place morality above self-love in their lives, but they fail to live up to this commitment, too often, on particular occasions. When they fail, doing wrong through weakness, they are morally responsible. Therefore, by Kant's theory, they choose (through *Willkür*) to act on the particular occasion contrary to their basic, life-governing maxim regarding self-love and morality (to which they are also committed through *Willkür*). This cannot be simply a momentary change of mind or heart – as if one could choose on everyday but Saturdays to treat morality as unconditionally and always prior to self-interest but then on Saturdays suddenly choose to make an exception of Saturdays, and then on Sundays again revert to the unconditional commitment, week after week. This is not a coherent picture, especially if the basic, life-governing maxim is viewed (as Kant presents it in *Religion*) as somehow representing a person's most fundamental choice over time.[39] The supposed "flip-flopper" or Saturdays-only immoralist is more

38 *G* 254–62 [4: 455–63].
39 *R* 61–73 [6: 39–53].

plausibly seen as having a morally impure will, which for Kant remains a form of "evil."[40] The morally weak, then, must be viewed as having two conflicting maxims at the time of weak-willed wrongdoing: a basic maxim to conform to morality's unconditional requirements and a shorter-term maxim reflecting an intention to indulge self-love on the particular occasion despite its conflict with morality and their own fundamental commitments. To preserve responsibility, in Kant's theory both must be choices through our power of choice (*Willkür*). There is, so to speak, a double conflict of will here: one freely chooses (through *Willkür*) to act contrary to both one's basic choice to prioritize morality (an act of *Willkür*) and to the inevitable demands of one's own rational legislative will (*Wille*). The virtuous person has neither kind of inner conflict.

This reconstruction raises a number of potentially troubling questions, which I can address only briefly here. First, we may ask, when someone does wrong through moral weakness, *what* is weak? We commonly say that the person's *will* was weak, but how does this play out in Kant's terms? The person's power of choice (*Willkür*) is presumed to have the freedom, or ability, to act independently of the causal force of inclinations. So it cannot be utterly weak in the sense "incapable of choosing" to do right. To conjecture that it is a "diminished capacity" to choose is to invoke empirical analogies that do not fit well with Kant's theory of agency. Nor is it necessary to invoke a "weak muscle" imagery to explain why pleading moral weakness tends to reduce blame but does not excuse wrongdoing. Those who lie or steal through moral weakness still have a basic commitment to morality and so at least have a conflict of will about the lying or stealing. They may be blamed less than those who behave similarly from an immoral policy that they whole-heartedly endorse, whether they follow the policy passionately or from cold, calculated self-interest.[41] The difference is not in what they *can* do, or can

40 *R* 53 [6: 30].

41 Kant implies that, other things equal, the degree of culpability for a wrongdoing is greater when the agent acts deliberately on a fixed bad policy than when giving in to strong inclinations (presumably against a moral disposition). He writes: "*Subjectively*, the degree to *which an action can be imputed (imputabilitas)* has to be assessed by the magnitude of the obstacles that had to be overcome. … [T]he less the natural obstacles ["of sensibility"] and the greater the obstacle from grounds of duty, so much the more is a transgression to be imputed (as culpable). – Hence the state of mind of the subject, whether he committed the deed in a state of agitation or with cool deliberation, makes a difference in imputation,

easily do, but in whether their bad act is a deviation, albeit a voluntary deviation, from their persisting, overall commitment to act well or rather a reflection of their bad fundamental commitments.

An alternative reconstruction would attribute the weakness to the agent's rational legislative will (*Wille*). At least one passage supports this interpretation, but it is apt to be more misleading than helpful.[42] The interpretation suggests that the explanation for wrongdoing through moral weakness is a deficiency in the agents' basic moral *capacities:* a crucial aspect of their practical reason (*Wille*) is sub-standard, unable to function effectively as it does in the virtuous. This invites us, when morally weak, to blame our equipment rather than our choices, which is hardly a Kantian idea. Moreover, attributing the weakness to the agent's rational legislative (*Wille*) still seems to reflect and reinforce the idea of moral conflict as a tug-of-war between two opposing forces, reason and inclination, leaving responsibility for choice out of the picture. We might say, speaking loosely, that practical reason (or *Wille)* was not as strong in the morally weak wrongdoers as in the virtuous, but, to avoid misunderstanding, we should avoid thinking of this as an explanation of failure due to defective *capacities.* Instead, we can view the expression as a loose, metaphorical way of describing someone who tends to yield to strong inclinations that are contrary to both their practical reason (*Wille*) and their fundamentally good will. Moral practical reason both *directs* and *predisposes* everyone to do what is right despite contrary inclinations, and in addition the morally weak (like the virtuous) have a *basic commitment* to do what is right. We might say that sometimes *all* of this "on the side" of moral practical reason (*Wille*) may not be "strong enough," as it were, to defeat the "other side" – strong contrary inclinations.[43] We would do better, however, to attribute moral weakness

which has results." *MM* 19–20 [6: 228]. Note that what mitigates culpability is not that the agent acts in an emotional state, which would not lessen the blame of those passionately committed to their immoral policies. It is the "state of agitation" that reveals a moral commitment in conflict with temptation that shows the weak wrongdoer's deed not as bad as the whole-hearted wrongdoer's otherwise similar deed.

42 *MM* 165 [6: 407].

43 It is noteworthy that in the troublesome passage where Kant seems to treat virtue as an aptitude "not as a property of choice (*Willkür*) but of will (*Wille*)", he characterizes the latter as "a faculty of desire that, *in adopting a rule*, also gives itself a universal law (my italics)." Here "will" (*Wille*) seems to refer to more than merely the aspect of practical reason that rationally *legislates* moral laws and *predisposes* even the worst moral agents to follow them.

simply to the *agents* themselves, or their particular choices, rather than to their moral capacities – *Wille* and *Willkür*. In a manner of speaking, they *will weakly* to do what is right, though they have adequate capacities to do otherwise.

How could this be so? Kant does not explain, and attempts to make sense of the idea must respect central points in Kant's action theory, especially that we act on maxims and that reason and inclination are not literally two opposing forces (phenomenal or noumenal). A first step might be to think of an individual's willing basic maxims on analogy with a governing body's exercising its power to adopt binding policies and regulations for its own members as well as others. Imagine it functions under a constitution that defines and limits its legitimate aims and activities, just as the fundamental moral laws of practical reason (*Wille*) impose limits on what basic maxims individuals can legitimately, or rightfully, adopt as their life-governing policies. The governing body may adopt rules and policies that are, in a sense, weak or strong, depending on their content. They are relatively weak insofar as they are vague and indeterminate about what is prohibited (or required) and relatively unspecific about the means of implementation. Weak policies are especially open to disagreements in interpretation, to self-deceptive special pleading, and to irregular and ineffective efforts at implementation. Similarly, in many cases moral weakness might be partially understood as a matter of individuals' incorporating the moral law into their basic life-governing maxim but in a somewhat vague and indeterminate form and without any specification about how to carry it out in the face of obstacles. This allows us to think that they still have the same fundamental moral commitment even though sometimes they willfully deviate from it. Their will of the moment is in conflict with their fundamental will to be moral, partly because the content of that will was weak, in a sense. But, assuming they were at least dimly aware of the conflict, this sort of weakness is not an excusing incapacity. Their basic maxim that gave precedence to morality over self-love was, in a sense, a weak maxim. It was not, however, the immoral basic maxim of a person who yields to the propensity to "impurity," for that would include acknowledged exceptions, as in "I will behave as morality prescribes except when this calls for significant sacrifices."

This sort of explanation would not, of course, account for morally weak deviations from explicit and definite basic maxims that even anticipate and specify how to guard against and manage contrary inclinations. Here we seem to need the idea that a person can will a fully

explicit and definite moral maxim, specific about means of implementation, but will it so weakly that deviations still make sense even though they are not excusable. The appropriate analogy might be a school board that, by full vote and repeated decisions of principle, adopts, reaffirms, and usually follows a "zero-tolerance" policy for expelling illegal drug users. When the mayor's son is caught with drugs, let us imagine, they give in to political pressure and make an unjustified exception. To make sense of the idea that they still, in a sense, hold the policy and yet knowingly deviate from it, we perhaps need to attribute to them imperfectly rational mental states and dispositions, regret, a guilty conscience, and a "will" to change. If so, we might say that they willed the zero-toleration policy, it remained their basic governing policy, but they failed to implement it in this particular case, without excuse. They willed the policy more weakly, we might say, than a different governing board that resisted all pressures to deviate from it. Some cases of moral weakness, perhaps, give evidence of a moral commitment, or will, that is weak in an analogous way. That is, our basic commitment over time was to an explicit and specific basic maxim to do what is right but, as imperfectly rational beings, we were capable of conflicts of will, choosing sometimes on particular occasions to deviate from deeper commitments that we never abandoned.

As Kant leads us to expect, such descriptions do not satisfy the criteria of empirical, scientific explanations, but I suspect that nevertheless they echo common moral thinking on these matters. And, from Kant's point of view, an advantage is that these ways of thinking do not treat reason, will, and inclinations as opposing forces of the same sort and they allow us to interpret morally significant acts as based on maxims. There remains, however, much that is puzzling and incomplete about this account. For example, some explanation is needed about why Kant allows (with Aristotle) that time and practice is required to develop virtue.[44] As strength of moral will, virtue does not consist in good habits and trained dispositions, but, like those, it cannot be acquired at will. What morally weak persons need is not, strictly, more power, ability, or habits of choice and feeling, but a better use of capacities they are presumed to have to adopt and implement a commitment to morality.

To conclude this section, I want to emphasize that my aim has been to sketch broadly Kantian ways of understanding some puzzling ideas in Kant's ethics and to suggest connections with familiar moral thoughts,

44 For example, *MM* 158–9 [6: 397].

but not to defend these ideas as correct or even strictly Kant's views. I turn now to some equally broad, and perhaps provocative, comparisons of Kantian ethics with "virtue ethics."

IV. A Brief Comparison: Kant's Ethics and "Virtue Ethics"

Finally, and all too briefly, let us take a look at some main themes commonly associated with virtue ethics to see how Kant's ethics compares. Here I will comment on four main points: (1.) the need for moral rules, (2.) sensitivity to particular contexts, (3.) morally good motivation, and (4.) the standard for right action.

IV.1. Are moral rules necessary?

Advocates of virtue ethics tend to deny or downplay any need for moral rules, and they especially object to strict, inflexible moral rules such as Kant's absolute prohibition on telling lies. First consider: Do we need rules of conduct at all? Where does Kant stand? In *law*, of course, he thinks we need public, authoritatively legislated and judicially enforced rules.[45] These should state explicitly what is prohibited and what is required, without incorporating *arbitrary* exceptions; but there is nothing in principle wrong, in Kant's view, with built-in complexities, including *well-grounded* qualifications. The laws, he held, should normally be enforced and obeyed rigorously, but there is one notable exception with implications far beyond what Kant himself envisioned, namely, we should not obey or (presumably) enforce a government's laws or orders if they require us to do anything "intrinsically immoral."[46] Regarding *social rules*, the shared moral codes taught and informally enforced by communities, Kant is remarkably unsupportive. Unlike Mill, who gave such rules a central role in his *Utilitarianism*, Kant relies almost entirely on general principles for law and for individual decision-making instead of informally enforced social codes. Contrary to his reputation, Kant was deeply opposed to moralistic meddling in others lives: our duty is to

45 *MM* 23–34 [6: 229–242], 89–120 [6: 311–351].
46 *MM* 136–7 [6: 371] and *R* 153n [6: 154n].

promote others' happiness, not to make them good.[47] Even judicial punishment, in his legal theory, is authorized only to prevent interferences with legitimate freedom, not to make wrongdoers suffer because they intrinsically deserve it.[48]

Laws and social codes, then, are not really the issue. Advocates of virtue ethics do not argue that *criminal and civil laws* are unnecessary, and Kant does not argue that *informally enforced social rules* are necessary. What remains controversial is whether *general principles for morally guiding individual moral decisions* are unnecessary. Critics charge that in Kant's theory these constitute a hierarchical system of ethical principles analogous to a legal system, and this legalistic ethics, they argue, serves no good purpose. A better alternative, they sometimes say, is to promote character development through the emulation of virtuous exemplars and the formation of good habits. Kant's position, recall, is that moral principles are sharply distinct from legal rules and, unlike legal rules, their purpose is not to serve as the basis for the mechanisms of public enforcement. The moral principles that he presents in his *Doctrine of Virtue* are a philosopher's attempt to articulate reasonable answers to recurrent moral questions that human beings face despite many differences in their circumstances. The principles are offered, with supporting arguments, in order to convince a reflective, critical audience that *doing what these principles prescribe* is necessary for a morally good life. Moreover, he proposes the principles as considerations that anyone would need to *respect as grounds for their decisions* in order to be fully virtuous. Behind this is Kant's thought that emulation, cultivated sentiments, and good habits are at best aids to virtuous living, not by themselves sufficient. Kant insisted, as in fact Aristotle did also, that judgments of practical reason are also essential; but Kant goes farther than Aristotle apparently did in claiming that the grounds on which practical judgment relies can be reliably and usefully articulated (at least by philosophers).

---

47 Evidence for this is assembled in Thomas E. Hill Jr., 1992, *Dignity and Practical Reason*, Ithaca: Cornell University Press, pp. 176–95.

48 Some evidence for this controversial thesis is presented in B. Sharon Byrd, 1989, "Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution", in: *Law and Philosophy,* vol. 8, no. 2, pp. 151–200; Donald E. Scheid, 1989, "Kant's Retributivism", in: *Ethics,* vol. 92, pp. 262–82 and in the chapter "Wrongdoing, Desert, and Punishment", in: Thomas E. Hill, 2002, *Human Welfare and Moral Worth: Kantian Perspectives*, Oxford: Oxford University Press, pp. 310–39.

IV.2. Are Kant's moral principles insensitive to relevant differences in contexts?

Advocates of virtue ethics sometimes suggest *any moral rule* is liable to oversimplify the complexities of particular situations, causing rule-followers to neglect morally relevant factors needed for correct judgment. This problem, if it is one, seems to be compounded by Kant because he endorsed rules, such as "Never tell a lie," that are simple, substantive, unqualified, and absolutely inflexible. What can we say on Kant's behalf? First, I think we must agree both that Kant endorsed such principles and that doing so was a mistake. We may question, however, how deep this problem goes in Kant's theory. Arguably, further reflection on the implications of Kant's basic moral framework does not actually support many of his extreme opinions about particular issues, such as lying, "unnatural" sex, and revolution.[49] We should also recall that Kant's main principles, beyond justice, for dealing with other people are "imperfect duties" concerned with moral attitudes and ends, for example, principles of gratitude, beneficence, respect, and friendship.[50] These, as Kant says, prescribe rather indefinite "maxims" rather than specific actions.[51] Whether even these are objectionably insensitive to context, as many of his "perfect duties" are, readers must judge for themselves.

Underlying the objection, however, may be a deeper issue. Are the grounds for moral judgment so relative to historical, cultural, and particular contexts that they can never be encapsulated in general principles? Particularists of various kinds answer "yes" and so reject moral principles, except as rough and potentially misleading heuristic devices. Most traditional moral philosophers, who were usually far more flexible about rules than Kant was, argue to the contrary that moral judgment presupposes that, at some level, there must be some common grounds for the legitimate variations in particular moral judgments. Bathing in lots of water may be fine in the monsoon season in India but wrong in the drought season in a desert community. There must be some reason, it seems, why the practice is wrong in one context and not the other, and the ultimate reason cannot be just that "that's what people say." Like utilitarians, perfectionists, intuitionists, and others, Kant tries to articulate and defend an account of ultimate and penultimate moral

49 *MM* 182–84 [6: 429–31], 178–80 [6: 424–26], and 95–100 [6: 318–25].
50 *MM* 198–218 [6: 448–74].
51 *MM* 150–56 [6: 385–94].

reasons. He expresses these in his fundamental principle (the Categorical Imperative) and some very general choice-guiding derivative principles, many of which are far from the specific "Do's" and "Don'ts" that critics of rules are most concerned to avoid. A number of Kant's specific principles, for example about sex and lying, are rightly rejected, but we do not need to embrace particularism of various sorts to avoid these.

IV.3. Does Kant ignore relevant features of good character and motivation?

Here again Kant's position is controversial and may be unsatisfactory, but misunderstanding and exaggeration often confuse the issue. Consider first Kant's views about good character. As we have seen, having a *good will* is the essential feature and to be fully virtuous we must have developed, in addition, *strength of will* to do what is right. We have the *particular virtues* of beneficence, gratitude, truthfulness, and so on, when we have a good and strong will to do what is right with respect to helping others, acknowledging our benefactors, communicating with others, and so on.[52] Besides these essential features, Kant acknowledges that there are other morally desirable traits of character (or perhaps personality) that we should try to cultivate, such as sympathetic feelings for others.[53]

Kant did not emphasize the need to develop affective sympathy, compassion, and other feelings, perhaps for two reasons. First, he saw British sentimentalism, his main competitor in moral theory, as false and dangerously seductive in its two respects: its explanation of the apparent authority of moral judgments and its idea of what makes acts morally commendable. Thus a major part of Kant's agenda was to emphasize the limitations of the role of feeling and emotion in moral judgment and motivation worthy of moral esteem. Second, he believed that wrongdoing and misguided moral judgment are to a great extent due to the way that passion, affects, and inclinations distort our moral judgment and provide self-deceptive excuses for making ourselves an exception to moral principles that we apply to others. His extreme and perhaps false empirical beliefs on these matters, arguably, bear considerable responsibility for his emphasis on the aspects of the moral life conceptually independent of affect and feeling. Arguably, however, Kant's basic moral

52 *MM* 157 [6: 395].
53 *MM* 205 [6: 457].

theory offers a more expansive role for feeling than Kant himself acknowledged. There are several points to note.

First, Kant did acknowledge some role for feelings in the moral life. Respect for the moral law, and through that for persons, has an affective aspect. In beings like us, he argues, recognition of moral requirements inevitably causes feelings of respect, akin in some ways to fear and awe.[54] Other moral feelings include a kind of satisfaction or "pleasure" at having done what is right and guilt and diminished self-esteem from having knowingly done wrong.[55] His thesis is not that we lack such feelings but, first, that they are a product of our moral awareness rather than its source and, second, that morally esteem-worthy motivation is not the same as doing what our feelings at the time prompt.[56]

Another point Kant conceded, but did not emphasize, is that we should cultivate positive other-regarding feelings to counter-act the negative feelings that tempt us to do wrong, thereby making it easier for us to do our duty. For someone committed always to do what is right, cultivating such positive feelings is among the means, or facilitating conditions, that should be taken. Kant grants that this is true even of the merely "aesthetic" or affective feelings for others, as distinct from the "practical" love and beneficence that consists of a commitment to promote others' well-being.[57]

But Kant's basic moral theory has the resources to say much more than this. Two further points in particular he *could* have added. First, from experience it seems that sympathy and compassion are important, and often necessary, to awaken our awareness of the needs of others and the way our choices may affect them. As feminist theorists rightly remind us, sensitivity to others and caring for them is needed to open our eyes to their needs and their suffering. Without these, callous, unfeeling moral agents, no matter how conscientiously determined to do what is right, will blunder repeatedly with potentially disastrous results. In short, affective sensibility provides essential information about the context that no one seriously committed to a moral life can allow him/herself to neglect.

A further point is this. Among the moral principles that Kant (and any reasonable moralist) endorses is a prescription to promote the

54 *CPrR* 62–75 [6: 71–89].
55 *MM* 15 [6: 221], 151 [6: 387], and 160 [6: 399–400].
56 *MM* 141–2 [6: 375–7].
57 *MM* 198–203 [6: 448–53].

happiness of others. In fact, the two primary ends Kant proposes in his *Doctrine of Virtue* are the moral and natural perfection of oneself and the happiness of others.[58] Kant was no expert about what specifically promotes happiness and what hinders it, but Kant's critics (among feminists and virtue theorists, for example) point to empirical facts that (arguably) Kant's basic moral theory requires us to take seriously into account. For example, whether you have genuine affection for your friend while visiting her in the hospital usually matters greatly to the quality of your relationship, to the happiness of your friend, and so to what feelings and attitudes you should cultivate and even to what you should do about visiting now, given how you feel. Cultivating our kindly feelings towards others and trying to transform our tendencies to indifference and feelings of hostility are required for virtually anyone who takes seriously the moral requirement, or desirability, of promoting others' happiness. Similar remarks could be made regarding other moral requirements that Kant acknowledges, for example, gratitude, friendship, and respect.

There is another point to consider. As Aristotleans remind us, failure to have the feelings morally appropriate to a given context is evidence that the agent is less than fully virtuous. This means, at least in part, that the agent does not have the sort of affective dispositions that we should all aim to acquire and that virtuous persons, through long practice by emulating exemplars, have succeeded in making habitual. In other words, according to the Aristotlean, everyone should aim to cultivate dispositions to feel in certain ways in various contexts even though success is not guaranteed because it depends on social setting, the availability of good mentors, and good luck. Having such dispositions to appropriate feelings, on this view, is a mark of a good and "virtuous" person in the sense that it amounts to success in achieving what everyone should strive for but not all can achieve, even with the best of efforts. Let us consider how this view compares with Kant's.

According to Kant, whether we have, or can develop, such traits depends as much on nature as on our good efforts. The mere possession of such traits, such as compassion, does not in itself make us *worthy of moral esteem*, but we can call them *good traits of character* to indicate that they are among the traits that, for moral reasons, we should strive to acquire. Even though Kant himself was inclined to minimize this point, arguably this follows from his basic moral theory together with a realistic understanding of human psychology. If we find someone lacking in the

58 *MM* 154–56 [6: 391–94].

moral feelings we come to expect of everyone, this does not *necessarily* mean that the person is morally bad or blameworthy but it may well be *symptomatic* of real moral failures. This is because, given human nature, if someone is *genuinely committed* to moral principles and *strives to become well suited* for a moral community, then typically that person will develop moral feelings as a natural consequence. Human beings are inescapably sensuous as well as rational. If our commitments, aims, and efforts are sincere, we are inevitably prone to feelings of disappointment when we fail, satisfaction when we succeed, and longings and worries when we are still striving. Similarly, if we genuinely respect others and "make their ends our own," as Kant directs, we are liable to indignation when oppressors abuse them, guilt when we ourselves mistreat them, delight when we see them succeed, and sadness when tragedy strikes them. Because of this, persistent emotional indifference to another's plight virtually always reflects a deeper failure of moral commitment. In Kant's view, moral feelings are the typical products and signs of moral commitment, though not its cause or what makes it worthy of esteem.

Now consider moral motivation. As we have seen, Kant thought that only acts from duty have "moral worth" in his special sense but, if one *has* a good will, acting from other motives in various situations can be appropriate and even commendable as exemplifying good character traits. A virtuous person, for example, may be commendable for caring for her children out of love, rather than duty, assuming that duty would move her if love should fade.

Another point also deserves mention. Critics sometimes ridicule the idea that virtuous people act justly and charitably "merely for the sake of duty" because they assume this means that they are focused on nothing but a lifeless, abstract principle. Understandably, they say, we should be concerned about the real people with whom we are interacting, not conformity to a philosopher's abstract idea. The problem here rests mostly on a misunderstanding, albeit one that Kant's own incautious remarks encourage. Expressions such as "respect for the moral law" and "acting for the sake of duty" call up images of submission to the commands of external authorities, divine and secular, but these images are inappropriate because utterly incompatible with Kant's thesis that the ultimate source of morality is the practical reason and autonomy of all moral agents. Your moral obligations stem from your own reason insofar as you reflect on human problems in a way that appropriately acknowledges the claims of others. In addition, we cannot "respect the moral law" if we disregard its core message. In the formulation most

directly concerned with the "matter" or content of the supreme moral principle, Kant says that we must treat all persons as ends in themselves, a status and unconditional value that demand that we treat them with justice, respect, and concern for their (permissible) personal ends as individuals. In acting from a commitment to morality, then, we are rightly focused on the dignity, needs, and aspirations of actual persons rather than merely an empty formula.

IV.4. What is the standard of right action?

Those associated with virtue ethics hold different views about the standard of right action, but some conceive of the right act as the act that a virtuous person would do in the situation.[59] Now most people, including Kant, would agree with this on one interpretation; but, as stated, the claim is ambiguous. We need to ask: Is the act right because a virtuous person would do it, or would a virtuous person do it because it is right? One cannot have it both ways. Kant held that, factual errors aside, a virtuous person does what is right because it is right. Then Kant offered criteria of rightness, attempting to draw relatively formal features from the concept of duty as a categorical imperative. Some virtue ethicists, apparently, hold that acts are right because they are what a virtuous person would do. Then they offer various suggestions about how to specify what a virtuous person would do. This position is coherent and understandably motivated by reaction to excessively legalistic moral theories, but nevertheless the position raises questions. For example, in cases where most clearly there is a right thing to do, are the standards of virtuous character by themselves determinate enough to identify the right act? They would be, of course, if virtuous character were defined in part as understanding and respecting the principles of justice, honesty, beneficence, gratitude, friendship, and so on. The difference between virtue ethics and Kantianism might be minimal if virtue ethics were to spell out these independent standards in certain ways. An extreme virtue ethics, however, would deny that there are such independent standards of right action. Arguably, without them, the idea of virtue becomes very thin, too insubstantial to guide our decisions or adequately explain morality.

---

59 For example: Hursthouse, 1999, especially pp. 17, 18, 28–31, 49–52.

Kant raises a related objection twice: when criticizing the divine command theory and when arguing that the supreme moral principle cannot be found by studying examples.[60] His point about the divine command theory is this: we could not derive our standards of rightness and goodness from the commands of a powerful deity because in order to *recognize the deity as worthy of obedience* we would *need to rely on our prior understanding* of rightness and goodness. Regarding empirical study of examples, he first notes that, for all we know, there may not be any actual examples of perfect virtue. His deeper point, however, is that trying to find the standards of right and virtue in examples is a fundamentally flawed procedure because *we must already have and apply such standards* in order to *identify* anyone *as an example.* Despite its far-reaching implications, Kant's basic idea in both cases is not radical but rather an idea implicit in ordinary moral thinking: that is, *to be virtuous* – and properly beneficent, grateful, honest, just, and so forth – *just is* in part to know and respect the *prior* reasons for acting rightly. And to respect these independent reasons is ultimately just to acknowledge fully the value of humanity in each person. To be sure, we should *do the things* that a fully virtuous person would and from the *same motives*, but *in moral theory* we should not try to derive our standards of rightness and goodness from alleged instances of perfect virtue, human or divine. And the morally best motive is respect for moral reasons, not a desire to emulate admired exemplars.

60 *G* 243 [4: 443], *G* 208–10 [4: 406–9].

Was Kant a Virtue Ethicist?

Robert N. Johnson

You might think a simple "No" would suffice as an answer. But there are features of Kant's ethics that appear to be strikingly similar to virtue oriented views, so striking that some Kantians themselves have argued that Kant's ethics in fact shares these features with virtue ethics. In what follows, I will argue against this view, though along the way I will acknowledge the features of Kant's view that make it appear more like a kind of virtue ethics than it really is.

My plan is to first set out the distinctive features of what is nowadays called "virtue ethics", those features that make it a genuine alternative to other normative theories. I then consider the features Kant's view might share in common with virtue ethics and the case for saying that it is, therefore, fundamentally the same sort of theory. I follow these two sections with an argument against this position. I want to warn you at the outset, however, that my argument itself will be quite unsurprising, since it is an argument that has been central to the way in which most philosophers have understood Kant's ethics. Any novelty I can claim here is in my account of what makes virtue ethics a genuine alternative to other normative theories, and my defense of this argument against those, in particular Barbara Herman, who have apparently found the argument unpersuasive.

I. The Form of Virtue Ethics

What makes a theory count as a kind of "virtue ethics"? Clearly it is not that the theory takes moral virtue seriously, or that it takes the virtues individually to be an important topic for ethical theory. That would make virtue ethics a trivial addition to an already theoretically sound group of alternatives in ethical theory. Virtue ethicists instead say they offer a genuinely new theoretical alternative, one fundamentally different from

deontological or teleological ethical theories.[1] Now one might assume that what makes it fundamentally different is that the theory orients itself around the moral virtues, rather than duty or valuable states of affairs.[2] Thus, instead of viewing virtue as merely an optimific disposition or disposition to right conduct – that is, as of interest only after the right and the good are understood – virtue ethics gives pride of place to the virtues and these in turn are to inform our understanding of the right and the good. That, at any rate, is how some philosophers have regarded virtue ethics.

But this assumption about what makes virtue ethics distinctive is misleading in a number of ways. First, it is not really virtue, but happiness or human flourishing, which has pride of place in these theories, especially those that claim ancient views such as Aristotle's as ancestors. Indeed, flourishing has to be foundational because it is the source of the list of virtues: The virtues are just those character traits that are integral to a flourishing life. Second, although human flourishing has pride of place in virtue ethics, it also is not really, or at least not precisely, what makes virtue ethics fundamentally different from the others. After all, utilitarianism also puts human flourishing at the center of its theory. What makes virtue ethics different from utilitarianism is in its understanding of the nature and role of human flourishing in a normative theory. For utilitarianism, flourishing is a valuable state to be furthered by our actions. For virtue ethics, flourishing is an activity, that of "doing well" as a human being. The latter conception of flourishing as an activity shows that virtue ethics has a radically different understanding of its role in normative theory. Hence, it is really the role human flourishing plays in the theory that makes virtue ethics distinctive.

More precisely, what makes theories collected under the heading of "virtue ethics" fundamentally different theoretical alternatives, if they are such, is, first and foremost, the form these theories take. What I mean by the "form" of a theory is how it arranges and argues for positions regarding the main topics of ethical theory. Those topics are the nature of right and wrong, of good and bad, and of what makes for a life worth

1 E.g., see John McDowell, 1997, "Virtue and Reason", in: Roger Crisp/Michael Slote (eds.), *Virtue Ethics*, Oxford: Oxford University Press, pp. 141–162.

2 E.g., "Certainly it is characteristic of modern virtue ethics that it puts primary emphasis on aretaic or virtue-centered concepts rather than deontic or obligation-centered concepts." Roger Crisp/Michael Slote, 1997 introduction to ibid., in: Crisp/Slote, 1997, pp. 1–25, p. 3.

living. Of course some philosophers, call them "pluralists", would argue that there are no systematic interconnections between the topics. But most hold that one of these topics is the primary and the other two secondary or derivative. While most ethical theories approach these topics by treating the good or the right as foundational, the alternative to these would build its ethics on what makes for a well-lived life. Hence, the virtues would only enter the picture theoretically once a conception is in place of what makes for such a life, or as I shall sometimes call it, an ideal of the person.

The claim that an ideal of the person, rather than duty or value, is the foundation of ethics is not, or at least not mainly, an epistemic claim. The central claim is not that necessarily one cannot know or have justified beliefs about which actions are right or which states of affairs are good without first knowing or having justified beliefs about what constitutes such an ideal. It could well be that we cannot know what a well-lived life is without first understanding moral duty and what is genuinely of value. But it may still be the case that what makes an action a moral duty and what makes something of value is its relationship to some ideal of the person. That may just be the only epistemic route to that grounding ideal. Nor is the claim necessarily conceptual or semantic, that is, a denial that concepts of this ideal are definable in deontic or value terms, or that statements about the ideal are true in virtue of obligation or value facts. Some might think that you can only define a well-lived life in terms of doing one's duty, others might well claim it goes the other way around, but that is not the same thing as arguing about which is foundational to theory. Moreover, the denial that right and value are foundational is also not primarily a *metaphysical* position, for instance, that an ideal of the person is "more real" or more capable of independent existence than the properties of right and wrong action or good and bad.

Virtue ethics is primarily a *normative* view and hence the central position has to do with what is basic *in a normative sense.* This means that the answer to traditional normative questions about the difference is between right and wrong, which states of affairs are good or bad, and what makes lives worthwhile, are, on the virtue ethical approach, to be arranged such that the last issue forms a basis for answers to the first two. Those actions are right and those things good that are related in the right ways to an ideal of the person, or so virtue ethics should say. Hence, you should act in the ways that characterize that ideal, and valuable states of affairs are just those that characteristically come from living that ideal. To be sure, pluralists will reject the idea that any of these notions should be

the basis of the others. But those virtue ethicists who think there should be some account of right action and of good states of affairs typically have in mind accounts that appeal to flourishing, as an ideal of how to live, as the foundation for right action and valuable states.

These features of the form of virtue ethics in fact show it to be, not *the* novel alternative to deontological or teleological ethics, but *one possible instance* of such an alternative. For were one to use the central claim that human flourishing is foundational to ethical theory to generate some account of right action, it would look something like this:

> VE: For all actions φ and all persons S, it is right (to be done, ethical, correct, etc.) for S to φ in C at *t* if and only if φing in C at *t* is or would be characteristic of a flourishing human life.[3]

Yet VE is an instance of a more general schema of this form:

> V: For all φ and all S, it is right (to be done, ethical, correct, etc.) for S to φ if and only if φing is (or would be) characteristic of P.

where "P" stands for some ideal of the person. To assume that VE represents the alternative to deontological or teleological approaches to ethics is to assume too much already, that the only ideal possible is one that would fit under the heading of *eudaimonia.* But it is, or at least can be, a substantive issue what sort of life is ideal and whether that ideal is a *eudaimonic* life. Hence, strictly speaking it is V, not VE, that is the alternative to deontological and teleological approaches to ethics. It is a bi-conditional according to which the left-hand side speaks of actions in deontic terms such as "right", and on the right-hand side speaks of the behavior characteristic of some ideal. The alternative theory says that we should generate an account of right action by appeal to the conduct characteristic of that ideal. For Aristotelians, P is the *eudaimonic* or flourishing person. But others, armed with a different ideal, would come up with a different set of actions that are right or to be done on V. What all such views share is the notion that an ideal, however that idea is spelled out, is the standard of correct conduct (and valuable outcomes), however that idea is spelled out. Such ideals (and note well that I leave it open *how* ideal they should be) are undoubtedly a central subject of ethical theory.

---

3 Hereafter, I omit the spatiotemporal markers for clarity's sake. The formulation is roughly what one finds in, for instance, Rosalind Hursthouse (see Rosalind Hursthouse, 1999, *On Virtue Ethics*, Oxford: Oxford University Press, p. 28); McDowell, 1997; and of course in Aristotle, 1993, *Nicomachean Ethics*, trans. by T. R. Irwin, Indianapolis: Hackett, 1105b6.

But it would be a mistake to assume that all such ideals must have something to do with the virtues or more generally with human flourishing, as normally understood in discussions of virtue ethics.

The form of V, then, is a critical component of virtue ethics, inasmuch as it is supposed to represent an alternative to deontological and teleological theories. But this is not yet sufficient to make it a genuine alternative. Utilitarianism and intuitionism, for instance, both could accept V, plug in their favored accounts of "right action" on the left hand side, and generate on the right hand side an ideal characterized by that sort of conduct. That, then, would be the theory's ideal of the person. The notion of such an ideal would simply be motivated by a prior notion of right action. So it is essential, if it is to be a genuine alternative, that virtue ethics makes the right hand side as far as possible *independent* of the left. That is, if virtue ethics is to be distinctive, the argument for the particular view of an ideal of the person (much less a view of human flourishing or happiness) must be innocent of premises presupposing some particular understanding of right action or good states of affairs. If it is not, then it would be simply building into the right hand side the elements of right action of which the ideal is supposed to be the source. The ideal of the person in the theory will be doing no work in generating answers about how to act.

The sense in which I mean that the ideal of the person "generates answers" is that it gives a rationale for particular lines of conduct; this will become clearer as we proceed. For now, notice that the other orientation of V, which takes the right as basic, "generates answers" about the ideal of the person at least in this sense: The question "What sort of ideal should I live up to?" gets answered constrained by first answering questions such as "What must I do?" and "What must I not do?" Whatever ideal conforms to the answers to those questions alone is ideal for those who take the left hand side as normatively basic.

II. The Case for Kantian Virtue Ethics

The account of the unique theoretical orientation of virtue ethics I've just offered consists of two features. First, it ties right action to an ideal of the person in the way that the bi-conditional in V does. Second, it generates the account of right action on the basis of an account of that ideal as far as possible by understanding the latter account *independently* of the notion of right action (or good states of affairs). Assuming this account is

correct, surprising as it may seem, one might naturally view Kant's own ethical theory as sharing both of these features with virtue ethics. Here is why.

The central idea is that the notion of a "good will" that inaugurates the *Grundlegung* and in particular the argument of *Grundlegung I*, offers just the form that supports this way of understanding Kant's ethics. Kant begins by specifying the nature of the good will apparently as if it characterizes an ideal of the person for rational agents such as ourselves, an ideal whose value is beyond measure and which is to be honored and respected. He then appears to use this as a basis on which to generate a rule of dutiful action, the Categorical Imperative. The idea, at least as it appears to be, is that it is characteristic of a good will to conform to the Categorical Imperative, and so this must be the fundamental principle of right or the nature of duty on the Kantian account. Thus, the assumption is that dutiful acts are those characteristic of a good will. And it is this assumption that has the form of V:

> KE: For all φ and S, it is S's duty to φ if and only φing is (or would be) conduct characteristic of a good will.

KE thus paints Kant's ethics as having precisely the same form as virtue ethics, the form of V, in which P is Kant's good will. Bringing to bear features of the argument of *Grundlegung* I, we are now also in a position to build a case for the second aspect of a virtue ethical theory, namely, that the disposition of a good will is the foundation of dutiful conduct in the way that the ideal is the foundation of right conduct in V. For the structure of that argument appears to support the view that Kant's ethics is a form of virtue ethics. Whether Kant's view is a kind of virtue ethical view will turn on the second crucial component of those views, the component that grounds the right in the ideal on the right hand side of V. Hence, it turns on whether the good will in Kant's ethics is in the right sense a foundation for the Categorical Imperative.

In what way might the good will – as an ideal on the model of ideals such as the Aristotelian *phronimos* – provide a foundation for the distinction between right and wrong conduct? If it is to play the foundational role the ideal does in virtue ethics, it must play the role of a final end. A final end in such theories is supposed to provide a grounding rationale for our behavior, a guide to deliberation that lays down a rule, principle or law for our behavior. The rationale, then, for Kant's ethics would be that the CI is laid down as a law for us because the value of the good will as a final end, an ideal of rational nature let us say, demands

this. The feature of the conduct violating the CI and making this constraint understandable is that the conduct is uncharacteristic of a good will.

Ends could be the source of practical laws in the following way: For Kant, a person's end is the reason she wills an action; as such, it is the ground of the laws guiding the will of that person. If it is not an end in itself, but an end only for some rational agents and not others, the law guiding the person's deliberation is only a hypothetical imperative, to take the means to that end or else abandon the end. Something that is an end in itself, by contrast, would be an end for everyone. Hence, it would have to be a reason for a different sort of law guiding deliberation, a universally valid categorical imperative.[4] If the good will, and more generally, the humanity in persons, is an end in itself, then it seems we have what we are looking for: Kant's ethics grounds the authority of the Categorical Imperative in a final end, the good will and (or) the humanity in persons (understood as including the capacity for a good will). The gulf between virtue ethicists and Kantians suddenly appears to be not nearly as deep as many may have assumed. The issue between the two views concerns only the content of an ideal for human beings, not whether some such ideal or other should be theoretically fundamental.

Before I assess this line of thought, there is one feature of KE I think that it is worth noting is a potential advantage for Kantian ethics over other doctrines that accept the form of V. The good will, as Kant conceives of this, is not necessarily a *virtuous will.* The lack of moral virtue, which Kant holds is the lack of moral strength of will to overcome obstacles to doing one's duty, is compatible with nevertheless having a good will.[5] Because of this, the ideal P as it appears in Kantian ethics is compatible with non-ideal character traits. That in turn allows us to ask a critical question for moral theory, namely, What would a good will characteristically do to improve his own character? The answer will then deliver, on the left hand side of the Kantian version of V, a duty to improve our own characters. I say this because, surprising as it may seem, any view that characterizes the ideal of the person as an ideal of *character* – as someone who by contrast *is* in possession of all of the virtues – will not be able to explain why we have a duty to improve our characters. It

4 [*G* 4: 427–8].

5 "Weakness in the use of one's understanding coupled with strength of one's emotions is only a *lack of virtue* and, as it were, something childish and weak, which can indeed coexist with the best will." [*MM* 6: 408].

will not be able to explain this because someone whose character is already completely virtuous – the very ideal on the right hand side of V in terms of which the theory is trying to generate a set of duties – would not characteristically engage in character-improving behavior.[6] This fact about Kant's theory, that his ideal of the good will is explicitly *not* cast in terms of virtues is, to be sure, an important difference between Kant's ethics and virtue ethics. But, as I have argued above, the role of the virtues in virtue ethics is not the source of its unique theoretical orientation. Indeed, if I am right, the dominant role of the virtues in such theories may even be a serious liability. As long as the ideal motivating the theory of right is conceived of as an ideal of character, the theory will not be able to explain duties of self-development.

I realize that in the eyes of many this fact alone will disqualify Kant's view as counting as a kind of virtue ethics; after all, virtue – either as a generalized disposition or as a set of particular traits – is not at all central to his theory. But if I am right, virtue is not really so central to theories called "virtue ethics" either. The critical move in these views is to draw our attention to the idea of the person, to its centrality and importance to our ethical concerns, and to challenge the notion that value or duty could occupy a place of that prominence in practical deliberation. That move does, to be sure, lead very naturally to a consideration of the enduring traits of character that are necessary to realize such a life. After all, so much of our own concerns with children are with instilling the sorts of habits and dispositions that will ensure that they have an opportunity to live a worthwhile life. Nonetheless, these considerations follow upon an original focus on the ideal of the person, an ideal about which Kant can be said to have an in-house disagreement with virtue ethicists.

III. Leaving Virtue Ethics Behind

My discussion so far has laid out the two features of virtue ethical theories that mark it out as a genuine alternative to other theories, and has offered what I take to be the reasons for thinking Kant's view shares those features. In this section, I explain why, nevertheless, Kant's position is wholly incompatible with a critical feature of virtue ethics, why, that is, Kant's orientation is as deontological as we always thought it was.

6 I argue this more fully in Robert Johnson, 2003, "Virtue and Right", in: *Ethics* vol. 113, pp. 810–834.

KE is itself not, so far as I can see, incompatible with Kant's views. So let us take as read that our duty is to conform to whatever principle that a good will embraces. So to this extent, we should acknowledge that Kant's ethics has the same form as virtue ethics. But this is only half the story. The critical question is whether the second element, the element that makes the right hand side of theories of the form of V basic, is also a part of Kant's views. I'll argue that this second element is quite incompatible with his ethics.

First, if the right hand side of KE is basic, then Kant must be able to characterize a good will independently of the principle of duty it supposedly generates. That is part of the idea of making the right hand side basic. But clearly this is wrong. There is no way of characterizing what a good will is, in Kant's moral philosophy, without deploying the idea of the Categorical Imperative. To be sure, appearances are deceiving: The order of presentation of ideas in the *Grundlegung* appears to set out the good will as something we already fully and clearly grasp. But in fact, this is only a rhetorical device, a device through which Kant can display how it is that we discover the fundamental principle of morality. Ordinary moral consciousness has a stark intuition of what it is that makes for a person of good will in his view, but even Kant would admit that that consciousness would not represent the Categorical Imperative so evidently as its guide. Nevertheless, Kant is confident that this is the principle each person deploys when he considers what he must do.[7]

Second, this order of discovery is most definitely not an order of normative priority. Kant is, after all, engaged in a *search for* the supreme principle of morality that Kant is undertaking in the first two sections of the *Grundlegung.*[8] In order to *discover* the fundamental principle of duty, Kant assumes that this principle will be one and the same as the principle of a good will. That is, he assumes something like KE. But this assumption is based on the idea that what makes a good will good is her disposition to deliberate and decide what to do by way of the fundamental principle of morality (whatever that principle turns out to be). The authority of the CI, the property in virtue of which it inspires respect in us, is not the fact that it is the principle of a good will. Quite the reverse. It is because of her embrace of the authority of the CI that her will enjoys the special value it possesses. Thus, whether or not the good

7 E.g., see [*G* 4: 403].
8 [*G* 4: 392].

will represents an ideal of some sort, it is not a fundamental notion out of which a theory of right action can be generated.

Third, the passages in which Kant argues that an end in itself – the humanity in persons – is a ground of the Categorical Imperative, as well as his discussion of the value of the good will, must one and all be understood against the backdrop of his general view of value. For instance, take perhaps the most well known passage that sets out the importance of the value of humanity:

> But suppose there were something the *existence of which in itself* has an absolute worth, something which as *an end in itself* could be a ground of determinate laws; then in it, and in it alone, would lie the ground of a possible categorical imperative, that is, of a practical law… Now I say that the human being and in general every rational being *exists* as an end in itself…[9]

To be sure, this passage appears to state that there is a unique value that is the basis of the moral law. Yet, as strong and clear as these words seem, they do not, in fact, say this. To have value, in Kant's view, is to be the object of rational willing.[10] Hence, for something to have absolute worth is for it to be, as it were, absolutely – or in every possible circumstance – the object of rational willing. So although the discovery of something with absolute worth is proof that there is a categorical imperative – an imperative that necessarily binds all rational agents – this does not show that the direction of normative priority in KE is from right to left. It does not show, that is, that it is the value of rational nature, the value of the good will, or any other value, that is the source of the categorical imperative. What it shows is that there is something that must be the object of any rational agent's will, and as such absolutely good. And if there is something that must be the object of any rational agent's will, whatever determines that will to make that thing its object would be a categorical imperative. That our wills are bound by a categorical imperative is why humanity and the good will have the value that they do, not the other way around.

Another way to put the point is this: It is not simply the *existence* of an end that lays down a law for deliberation and decision; it is the *willing* of an end that lays down a law.[11] Being an end is nothing over and above being the object of rational willing. And so being an end in itself is

9 [*G* 4: 428].
10 [*CPrR* 5: 58–65].
11 [*MM* 6: 385].

nothing over and above being *in itself* an object of rational willing. The reason *willing* an end gives one a law is because rational willing is, in Kant's view, autonomous. A rational will must, by virtue of being a rational will, be committed to universal law. To be sure, Kant's position is astonishing and controversial. But it is at the very heart of his ethical thought. To think that there must be some value to be a ground for conforming one's will to moral law is to think that a rational will is heteronomous.

These considerations against a virtue ethical reading of KE are not *recherché*. They are elements of a standard understanding of Kant's ethics. Yet some Kantians have argued that, nevertheless, this standard understanding is entirely wrong, that Kant's ethics shares the unique theoretical orientation of virtue ethics I set out above. In particular, Barbara Herman has argued that Kant's view is not a "traditional deontology" in which there is an "absolute priority of the right".[12] Deontology, she argues, has no conception of value at its foundation. And that means that there is nothing in the theory to offer a "rationale" that

> renders moral action intelligible to the agent making possible the reasoned integration of morality into one's system of ends [and] introduces a framework for reasoned deliberation necessary to the stable resolution of morally complex situations. A grounding conception of value could provide this rationale by offering an explanation of the wrong- or right-making characteristics of action that renders moral requirements intelligible in a way that is then able to guide deliberation… One way of putting the question to Kantian ethics is to ask whether the unconditioned good – the good will – can play the didactic role of a (the) final end.[13]

Herman understands the priority of the good as its playing a "didactic" role of a final end, an end that can guide deliberation and allow moral constraint to make sense to the agent in terms of her own system of ends. She is clearly not speaking of "the good" as a state of affairs to be produced by us, and so is not suggesting that Kant's project is parallel to consequentialism. She is not, in particular, supposing that conforming to the CI is required of us because in so doing we will produce a good will in ourselves. She is thinking of the good will as occupying the selfsame position in a theory that virtue ethicists think of *eudaimonia* or human flourishing as holding. In her view "rational nature is … the regulative

12 Barbara Herman, 1993, "Leaving Deontology Behind", in: Barbara Herman, *The Practice of Moral Judgment*, Cambridge MA: Harvard University Press, pp. 208–240, at p. 210.

13 Ibid. p. 216.

and unconditioned *end* of willing – that is, a final end, an end-in-itself".[14] What makes this possible is that rational nature is "autonomous, having the capacity to be its own original source of reasons." This makes rational nature "a unique kind of value", the kind of value that can play the didactic roles Herman, and she supposes, Kant, set out for it.[15] The basic idea is that you need to be able to say why it is good to conform to some rule if that rule is to guide your deliberations and decisions. But, according to deontological views, you are not able to say this with regard to moral rules. Luckily, in her view, the whole account of duty in Kant's ethics is in fact based on a conception of value, and that is the conception of the good will (and the related notion of humanity in persons).

Herman does not cast her view in as clear a language as one might like. She says that "we can understand the formal requirements of practical reason as a conception of value"[16] and that "we cannot understand what practical reason is without understanding it as a conception of value"[17]; the "purely formal principles of rationality can be expressions of … conceptions of value".[18] But neither a principle of rationality nor the faculty of practical reason is a conception or expression of anything. We *could* understand her point as being that a principle, in binding our will, does so by stating what it would be good to will, or else that rational principles are standards of valuable behavior, standards of what it would be good to do. But if this were her meaning, then there would be no reason to suppose that this is not a deontology. It is just to restate Kant's view that the good is the object of a rationally determined will, a will determined by principles of practical reason. Those principles, or at least the categorical imperative, cannot themselves be, in turn, based on some value, since embracing those principles is supposed to compose what willing anything at all consists in.

In fact, Herman anyway explicitly rejects this way of understanding her point.[19] She wants to know why conforming to the principles of rationality is good, how rationality itself is a value.[20] I can only think that this must just be the question, What reason do we have to be rational? Hence, if Kant's view holds that we must conform to the Categorical

14 Ibid.
15 Ibid. p. 238.
16 Ibid. p. 239.
17 Ibid.
18 Ibid. p. 236.
19 Ibid. p. 215.
20 Ibid. p. 213.

Imperative because it is a demand of rationality, we now need to know what reason there is to be rational – what good is it to be rational. Here, however, we have to distinguish between the faculty of rationality and the various principles that characterize rational willing. It may be that, as for instance John Broome has argued, there is reason to possess the faculty of rationality – i.e., that set of dispositions to conform to various principles of rationality – but no reason to satisfy any individual principle of rationality.[21] It may be so because while it may be in our interest to be rational, this does not show, and in Broome's view there seems to be no way to show, that there is a reason to satisfying any individual principle. They are two different questions.

I think that although Herman often states her question as what rationale there can be for moral constraints, I believe she is not looking for a reason to conform to the Categorical Imperative – an individual principle of rationality. What Herman wants is a reason to possess the faculty – characterized, to be sure, as Kant uniquely characterizes it – of rationality, since she speaks of the set of rational principles as a conception of the good. She wants to know what good there is in possessing *that* faculty of rationality, characterized as encompassing dispositions to follow moral as well as instrumental principles. But, unlike Broome, she does not want an answer in terms of our interests. In particular, if the rationale for conforming to moral constraints is simply that we must be rational – possess the disposition to conform to such constraints as well as the other dispositions – then we need some way to understand being rational as intrinsically valuable. It must, as *eudaimonia* does in the Aristotelian theory, all by itself end questions about the point or purpose of action. Unless we have this endpoint of deliberation, she thinks, we are left unable to resolve deliberative conflicts, such as when morality asks us to act in ways that are deeply troubling to us. What rationale is there for a moral constraint that asks us to give up things of such great importance that to do so will seemingly damage ourselves or those we care for? We must be able to see it as in some way the realization of a certain kind of intrinsically valuable identity as a rational agent.

The crucial move in this story is to appeal to rationality as a kind of intrinsic value. But Herman's ending the account at this point cannot be squared with Kant's overall theory of value. Although she points out that things come to have value in Kant's view by being the objects of rational

21 John Broome, 2005, "Does rationality give us reasons?", in: *Philosophical Issues*, vol. 15, pp. 321–37.

willing, she apparently does not regard this as true also of things that are intrinsically and unconditionally good. But this is a mistake. To be intrinsically and unconditionally good is also to be the object of rational willing. There can be no *independently existing* value in Kant's value theory, not even humanity and the good will.[22] Humanity and the good will are ends in the sense that they are to be respected in the ways Kant discusses. But it is not that Kant is pointing out that these are a kind of non-natural gem existing in the world whose value demands certain attitudes and behavior on our part. To have value is nothing over and above being the object of these sorts of rational demands.

There is an important reason why Kant's theory cannot be read in the way Herman wants to read it. As she points out, of course, that theory denies that there is a *nonmoral* rationale for moral constraint. Hence, it denies that moral action needs "the reasoned integration of morality into one's system of ends", *if* those ends are understood as *nonmoral* ends. That is not the failure Herman is concerned with. But the whole point of a deontological view is to argue that moral agents *already perceive or embrace* moral principles. It thus follows that if you are rational enough to know you have a duty, you already will to conform to it. That is why there is nothing in Kant's ethical philosophy to answer the question "Why be moral?" Any minimally rational agent in his view *already* possesses a will that, by its very constitution, wills the moral law. That distinguishes it from deontological views that claim that any rational agent must already *perceive* the moral order.

It is thus central to Kant's position that moral philosophy does *not* address questions about the rationale for moral constraint, even if it must address why, if there are moral constraints at all, every rational agent is already in touch with them. This is what makes sense of Kant's position that, unlike those such as Hobbes and Hume who came before him, there is no need to set out a reason to conform to the Categorical Imperative. It is not the fact that those theories try to tell us why, in terms of our *subjective contingent* ends, we should conform to moral restrictions that he rejected them and regarded them as "heteronomous" theories.[23] What makes them heteronomous is that their very conception of the rational agent is of an agent whose will does not already commit himself to the

22 I argue this point in some detail in: "Value and Autonomy in Kantian Ethics", in: *Oxford Studies in Metaethics,* vol. 2, R. Shafer-Landau, ed., Oxford: Oxford University Press 2007, pp. 133–148.

23 [*G* 4: 441–445].

moral law. If that is right, then wanting to give a rationale for moral constraint to rational agents makes less sense than giving Yankees fans a rationale for wanting the Yankees to win.

It might now be thought that if, as I say, a central feature of Kant's ethics is a conception of persons as already committed to the moral law, then his theory, after all, *does* ground itself in an ideal of the person. It is an ideal of the rational agent as already committed to the moral law. But, first, the conception of rational agency here does not play the role in the theory that it would have to play in order to provide a rationale for moral constraint. The role Kant's conception of agency plays, instead, as I explain above, makes such a rationale otiose. Second, Kant's conception of rational agents is itself wholly informed by a prior understanding of the moral law. As such, it obviously cannot be then used to provide a rationale for the moral law.

IV. Conclusion

I said at the beginning that my point would not be surprising. Although there are theories that provide a final end to play the didactic role in deliberation that Herman sets out, I think Herman is quite wrong in arguing that Kant's could be one of them. Kant does not offer anything to make morality intelligible to someone whose ends are not already arranged around, and whose deliberations are not already informed by, the moral law.

The Priority of the Right in Kant's Ethics

Jörg Schroth

For Kant the right has priority over the good – at least that is how he is usually interpreted. The origin and Kant's explicit statement of the doctrine of the priority of the right is in the second chapter of "The Analytic of Practical Reason" in the *Critique of Practical Reason:* "On the concept of an object of pure practical reason", where Kant writes that

> the concept of good and evil must not be determined before the moral law (for which, as it would seem, this concept would have to be made the basis) but only (as was done here) after it and by means of it [*CPrR* 5: 62 f.].

Yet, Allen Wood has recently claimed that Kant accords the good priority over the right because in *The Doctrine of Virtue* he grounds our duties in the promotion of ends:

> Because Kant bases all specific ethical duties on our virtuous commitment to ends, within the system of ethical duties he grounds the duty to act in certain ways exclusively on the promotion of ends. In the language of twentieth-century Anglophone ethical theory, this means that, within the system of duties, he holds to the priority of the 'good' over the 'right', and is therefore a 'consequentialist' rather than a 'deontologist' in the main senses those terms now have for moral philosophers. But, of course, the *fundamental principle* on which Kant grounds ethics is not consequentialist. This points to the importance of distinguishing the *fundamental principle* of an ethical theory from the *style of reasoning* it recommends in ordinary deliberation. We may (as Kant does) advocate consequentialist reasoning in much moral deliberation without accepting a consequentialist foundation for morality.[1]

Did Kant then shift from the priority of the right in the *Groundwork* and the second *Critique* to the priority of the good in the *Doctrine of Virtue?* If not, does the fault lie in Wood's interpretation of the *Doctrine of Virtue* or in the common reading of the earlier works: Is the right prior to the good in the *Doctrine of Virtue* too or is the good prior to the right even in the *Groundwork* and the second *Critique?* How does this affect the

1 Allen W. Wood, 2002, "The Final Form of Kant's Practical Philosophy", in: Mark Timmons (ed.), *Kant's Metaphysics of Morals. Interpretative Essays*, Oxford: Oxford University Press, pp. 1–21, at p. 14.

traditional deontological and the recent teleological and consequentialist readings of Kant as well as his relation to virtue ethics? In the following paper I will try to answer these questions.

I.

I begin with Rawls's famous account of the priority relation between the right and the good. Rawls first mentions the priority of the right in *A Theory of Justice.*[2] There the sentences which clearly express what the priority of the right amounts to are the following:

- Conceptions of the good have to conform to what the principles of justice require.
- The principles of right, and so of justice, put limits on which satisfactions have value; they impose restrictions on what are reasonable conceptions of one's good.
- Men's desires and aspirations are restricted from the outset by the principles of justice which specify the boundaries that men's systems of ends must respect.
- Certain initial bounds are placed upon what is good and what forms of character are morally worthy, and so upon what kinds of persons men should be.

What these sentences come down to is variously expressed as:

- The good is not defined independently of the right.
- The right determines permissible and impermissible conceptions of the good.
- The right constrains the good.

Accordingly, the priority of the good means:

- The good is defined independently of the right.
- There are no impermissible conceptions of the good.
- The good is not constrained by the right.

What precisely does it mean that the good is defined (not) independently of the right?[3] As a fairly uncontroversial point I take it that some goods,

2 John Rawls, 1999, *A Theory of Justice. Revised Edition*, Cambridge, Mass.: Harvard University Press, p. 28.

3 According to Rawls the definition of the good independently of the right "means two things. First, the theory accounts for our considered judgments as to which

e.g., pleasure, preference satisfaction, etc., are non-moral goods. Non-moral goods are defined independently of the right because they can be established as goods without recourse to conceptions of the right. Hedonism, e.g., which tries to establish that pleasure is the sole intrinsic good is a theory of the non-moral good or, which amounts to the same, a non-moral theory of the good: arguments for and against hedonism do not turn on moral considerations, and theories of the right cannot strengthen or weaken the case for hedonism. Most non-moral goods are personal goods, i.e., goods which promote the well-being of individuals, like, e.g., the goods on Griffin's[4] list of prudential goods that are valuable for every human life and components of a good life: accomplishment, the components of human existence (which include among other things freedom from pain and anxiety, autonomy and liberty), understanding, enjoyment, deep personal relations. These non-moral goods must be distinguished from moral goods (which, like justice and fairness, are often structural goods as opposed to the non-moral personal goods).[5] Moral goods cannot be established as goods without recourse to conceptions of the right and hence cannot be defined independently of the right.[6] With

---

things are good (our judgments of value) as a separate class of judgments intuitively distinguishable by common sense, [..]. Second, the theory enables one to judge the goodness of things without referring to what is right" (Rawls, 1999, p. 22).

4 James Griffin, 1986, *Well-being. Its Meaning, Measurement, and Moral Importance*, Oxford: Clarendon Press, p. 67 f.

5 I cannot argue here for the distinction between the moral and the non-moral good. It is an intuitively plausible distinction which underlies most accounts of utilitarianism and consequentialism. See, e.g., how Darwall begins his characterization of consequentialism: "Consequentialism begins with the idea that there are values that are prior to morality. Even if there were no moral right and wrong, some things would still be good and others bad. The pain and suffering caused by a cataclysmic earthquake, for example, are bad things, regardless of any relation to vice or misconduct. They are bad things *to happen,* bad states of the world. […] Both the agent-relative and the agent-neutral disvalue of pain are independent of morality, […] so both are called 'nonmoral'. […] A fundamental tenet on which all consequentialist moral theories agree is that the moral rightness and wrongness of acts is determined by the *nonmoral value* of *relevant consequences*" (Stephen Darwall, 2003, "Introduction", in: Stephen Darwall (ed.), *Consequentialism*, Oxford: Blackwell, pp. 1–7, at p. 1 f.).

6 This point is not uncontroversial. One might hold that not every moral consideration involves a conception of the right and that even though moral goods are not defined independently of moral considerations, they may still be defined independently of conceptions of the right. I cannot discuss this matter here and simply assume that moral goods cannot be defined independently of a

this distinction between moral and non-moral goods in place we can define what it means that an ethical theory defines the good independently or not independently of the right. A necessary condition for the good to be defined independently of the right is that the theory of the good does not include moral goods. If the exclusion of moral goods is also regarded as a sufficient condition we get the following *first account of the definition of the good independently/not independently of the right.*

- The good is defined independently of the right if the theory of the good includes only non-moral goods.
- The good is not defined independently of the right if the theory of the good includes moral goods (besides non-moral goods).

A few examples might be helpful. First, take W. D. Ross's theory of the good which consists of four things that are intrinsically good: "virtue, pleasure, the allocation of pleasure to the virtuous, and knowledge (and in a less degree right opinion)".[7] Since virtue[8] and the allocation of pleasure to the virtuous[9], are moral goods Ross does not define the good independently of the right. Natural law theories, on the other hand, define the good independently of the right because their list of goods consists only of non-moral goods: life, knowledge, play, aesthetic experience, sociability (friendship), practical reasonableness and religion.[10] Classical hedonistic utilitarianism also defines the good inde-

conception of the right. At least as regards distributive justice this accords with Rawls's view that "if the distribution of goods is also counted as a good, perhaps a higher order one, and the theory directs us to produce the most good (including the good of distribution among others), we no longer have a teleological view in the classical sense. The problem of distribution falls under the concept of right as one intuitively understands it, and so the theory lacks an independent definition of the good" (Rawls, 1999, p. 22).

7 W. D. Ross, 1930, *The Right and the Good*, Oxford: Oxford University Press, p. 140.

8 "The first thing for which I would claim that it is intrinsically good is virtuous disposition and action, i.e., action, or disposition to act, from any one of certain motives, of which at all events the most notable are the desire to do one's duty, the desire to bring into being something that is good, and the desire to give pleasure or save pain to others" (Ross 1930, p. 134).

9 "Thus the pleasures of which we can say without doubt that they are good are […] the pleasures of moral beings that are deserved and are either realizations of good moral dispositions or realizations of neutral capacities (such as the pleasures of the senses)" (Ross 1930, p. 138).

10 This list is taken from John Finnis (1983, *Fundamentals of Ethics*, Oxford: Oxford University Press, 50 f.). Natural law theorists do not agree on a list of objective goods. Alfonso Gómez-Lobo (2002, *Morality and the Human Goods. An*

pendently of the right and differs in this respect from consequentialist theories which include distributive justice among the goods to be maximized. But granting this distinction between defining the good independently and not independently of the right, why should it be appropriate to call this a difference in the *priority* of the right or the good? In what sense of "priority" does the good have priority if it is defined independently of the right? We may find an explanation when we remember that Rawls's talk of the priority of the right/good in *A Theory of Justice* follows his now famous classification of ethical theories in terms of how they define and connect the right and the good. He assumes that by this criterion all ethical theories can be divided into two classes:

A theory is teleological if and only if

– the good is defined independently of the right, *and*
– the right is defined as that which maximizes the good (Rawls 1999, pp. 21 f.).

Deontological theories are defined as non-teleological theories, hence as theories where

– the good is not defined independently of the right, *or*
– the right is not defined as that which maximizes the good.

Usually these definitions are taken to correspond to the priority of the good and the priority of the right, respectively. The idea behind the definition of the good independently of the right in teleological theories is that these theories *first* define the good and *then* define the right as the maximization of the previously defined good. The good is thus defined literally prior to the right. Because it is defined prior to the right, the good cannot include moral goods (since moral goods cannot be specified without reference to the right). Therefore, if the good is defined prior to (i.e., before) the right it must be defined independently of the right in the sense defined above.

This sense of the priority of the good in teleological theories is untenable. To begin with, teleological theories do not need to define the good first and the right afterwards. They might as well begin with a definition of the right, namely that the right consists in the maximization

---

Introduction to Natural Law Ethics, Georgetown: Georgetown University Press, pp. 6–25), e.g., defends the following list: life, family, friendship, work and play, experience of beauty, knowledge and integrity.

of the good, and then define the good.[11] Having started with this conception of the right, one needs to find out which are the intrinsically good things that ought to be maximized. Whether one begins with a theory of the good or a theory of the right is neither here nor there: a theory of the non-moral good is part of any plausible moral theory. It is logically independent of theories of the right and may be established before or after the theory of the right. Whether a moral theory has room for moral goods besides non-moral goods does not depend on the order of justification of the theories of the good and the right. Rather, it depends on what kind of theory of the right the moral theory consists of. If, like in utilitarianism, the theory of the right consists of one principle only, viz. the maximization principle, the moral theory cannot include any moral goods because the inclusion of moral goods would make the theory circular. This point has been made by Frankena who claimed that in utilitarianism the moral quality of actions can only depend on the *non-moral* value of what they bring about because: "For the moral quality or value of something to depend on the moral value of whatever it promotes would be circular."[12] There is indeed a circle if one holds that the moral value of an outcome depends on the right and then claims that the right depends on the moral value of the outcome. If the moral good depends on the right (i.e., if what counts as a moral good depends on the conception of the right) and if the right consists *only* in the requirement to maximize the *moral* good, then it is impossible to know what it is that one ought to maximize. If the moral good is not identifiable independently of the right and the right, in turn, consists *only* in the requirement to maximize the moral good, the theory is circular and cannot identify justice or fairness or anything else as a moral good. If, instead, the right consists only in the requirement to maximize the *non-moral* good one knows what to maximize (pleasure, for example). But of course, nothing in the theory of the non-moral good compels us to accept only one moral principle in our theory of the right. If we add further moral principles or subscribe to an altogether different set of moral

11 This point is also made in Roger Crisp's entry on deontology in the *Oxford Companion to Philosophy:* "[Utilitarianism] can suggest that the right is indeed prior to the good, in the sense that utilitarians can state that it is right to maximize the good, whatever the good turns out to be" (Roger Crisp, 1995, "Deontological Ethics" in: Ted Honderich (ed.), *The Oxford Companion to Philosophy*, Oxford: Oxford University Press, pp. 187 f., at p. 188).

12 William K. Frankena, 1973, *Ethics, Second Edition*, Englewood Cliffs, N. J.: Prentice-Hall, p. 14.

principles which does not include the maximization principle we can without circularity add moral goods to the theory of the good.

To sum up, whether a moral theory's theory of the good consists only of non-moral goods or includes moral goods as well, i.e., whether the good is defined independently or not independently of the right, depends entirely on the theory's principle(s) of the right. Defining the good prior to and therefore independently of the right amounts to nothing more than defining the non-moral good. The definition of the non-moral good is part of any plausible moral theory and decides nothing as to whether there are moral goods besides the non-moral goods. Depending on which theory of the right we subscribe to, we may or may not include moral goods. We may therefore draw the following conclusions: if the priority of the good is understood as defining the good before defining the right, the priority of the good does not entail that the good is defined independently of the right. If the priority of the good is understood as defining the good independently of the right, it remains unclear which sense of "priority" could warrant to refer to the independent definition of the good as the priority of the good. One might even say that, if at all, the right has priority over the good because it is the theory of the right that determines whether a moral theory contains moral goods or only non-moral goods.

What about defining the priority of the right/good in terms of permissible and impermissible conceptions of the good? After all, Rawls's main concern is not whether a theory includes moral goods or not but rather whether a theory distinguishes between permissible and impermissible conceptions of the good. Even in a theory that includes only non-moral goods the selection of which things are to be included into the theory of the good might depend on the theory of the right if the theory of the right excludes some non-moral (alleged) goods as impermissible conceptions of the good. To accommodate this, the above-mentioned account should be amended to the following *second account of the definition of the good independently/not independently of the right:*

- The good is defined independently of the right if the theory of the good includes only non-moral goods *and* the selection of these non-moral goods does not depend on the theory of the right, i.e., the theory of the good includes only non-moral goods and there are no impermissible conceptions of the good.
- The good is not defined independently of the right if the theory of the good includes moral goods, *or* (regardless whether the theory includes

> moral goods or not) if the selection of the non-moral goods to be included in the theory of the good is constrained by the theory of the right insofar as some goods are excluded from the list because they are incompatible with the theory of the right. In short, the theory of the good is not defined independently of the right if it includes moral goods or if there are impermissible conceptions of the good.

For Rawls, the paradigm of a theory with no impermissible conceptions of the good is utilitarianism: In calculating the best outcome *every* pleasure or preference must be taken into account. No pleasures or preferences are ruled out in principle from this calculation. Even the most horrible preferences must be impartially balanced (according to their strength) against other preferences. Whether a horrible preference (like, e.g., the desire to rape someone) may be fulfilled depends on the outcome of the calculation which is determined by the details of each situation and not by considerations of principle.[13] If, as a result of this calculation, a preference is not allowed to be fulfilled, this prohibition applies only to the case at hand with no implications for other cases: The preference had to give way to other stronger preferences, but has not been ruled out in principle from consideration and has to be considered again in each newly arising case. In Rawls's theory of justice as fairness, on the other hand, there are some conceptions of the good which are ruled out by the principles of justice as having no value and being wrong in itself: The "pleasure in discriminating against one another, in subjecting others to a lesser liberty as a means of enhancing their self-respect" (Rawls 1999, p. 27) is an impermissible conception of the good that need not (and must not) be considered in moral deliberations. It is impermissible because the first principle of justice demands that "each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others" (Rawls 1999, p. 53) and experiencing the pleasure of discriminating against others necessarily involves a violation of this principle.

This difference between utilitarianism and Rawls's theory seems clear enough. But is it really a difference in kind, a difference between the priority of the good and the priority of the right, a difference that divides ethical theories into two kinds, teleological ones and deontological ones?

---

13 Of course, we may know in virtue of empirical generalizations that some preferences will nearly always be outweighed by other preferences and therefore need not be considered each time anew. But this exclusion from consideration is justified by empirical, not moral, considerations.

That may well be doubted. To see this, let us first define an impermissible conception of the good as a conception of the (moral or non-moral) good which the theory of the right excludes from being allowed to pursue and which need not (and must not) be considered in any deliberations about what is the right thing to do. Preference utilitarianism, which demands the maximization of overall preference satisfaction, places no restrictions as regards the content of preferences, and everything that can be taken as someone's preference must be taken into account. Even someone's preference for equal distribution of goods, which runs contrary to the utilitarian goal of maximization, must be considered. However, there is a variant of preference utilitarianism which defines the right in terms of *personal* preferences only and treats *external* preferences as impermissible conceptions of the good. Does it make sense to say that in this theory the right has priority over the good whereas in an all-out preference utilitarianism the good has priority over the right? Or take Harsanyi's requirement that anti-social preferences be excluded from consideration.[14] Does Harsanyi accord the right priority over the good? The same question might be asked of Brink's[15] objective utilitarianism. And further, should these theories then be counted as deontological since the priority of the right is typical of deontological theories? Even classical utilitarianism has impermissible conceptions of the good, e.g., equal distribution and retributive punishment (understood as intrinsic, not as instrumental goods). Of course, depending on the theory of the right, moral theories have more or less impermissible conceptions of the good. Maybe there is only one theory with no impermissible conceptions of the good, namely an all-out preference utilitarianism. But to hold that in every theory except preference utilitarianism the right has priority over the good was surely not the intended meaning of the priority of the right. However, one could indeed say that in every moral theory, except virtue ethics, the right has priority over the good because it is the theory of the right that determines which, if any, conceptions of the good are

14 "[S]ome preferences, which may very well be their 'true' preferences under my definition, must be altogether excluded from our social-utility function. In particular, we must exclude all clearly antisocial preferences, such as sadism, envy, resentment, and malice" (John C. Harsanyi, 1977, "Morality and the Theory of Rational Behaviour", in: Amartya K. Sen and Bernard Williams (eds.), *Utilitarianism and beyond*, Cambridge: Cambridge University Press 1982, pp. 39–62, at p. 56).

15 David O. Brink, 1989, *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.

impermissible. As with the other two conceptions of the priority relation between the right and the good, this conception cannot give a plausible account of the *priority* of the good. Furthermore, the equation of the priority of the right with the existence of impermissible conceptions of the good does not indicate a useful distinction between two kinds of moral theory. It is rather a matter of degree how many impermissible conceptions of the good a theory has. A very specific theory of the right (which, like Rawls's theory, for example, refers to equal basic liberties) has more impermissible conceptions of the good than a less specific theory of the right (which refers only to pleasure).

Although the three conceptions of the priority relation between the right and the good as considered so far are not very helpful, we may ask whether, according to these conceptions, Kant gives the right priority over the good. As regards Kant's theory of the good I rely on its Rawlsian interpretation. Rawls[16] distinguishes six conceptions of the good in Kant's theory:

(1) Our conception of happiness.
(2) The fulfilment of true human needs.[17]
(3) The fulfilment of permissible ends, i.e., ends that respect the limits of the moral law.
(4) The good will.
(5) The ideal of a realm of ends, i.e., the good as the object of the moral law.
(6) The complete good.

The first conception of the good is given by "unrestricted empirical practical reason" and must satisfy only "various principles of rational deliberation that characterize the hypothetical imperative". It is the only conception of the good that is "entirely independent of the moral law, since it is the rational without restriction" (Rawls 2000, p. 231). The second conception of the good is designed to ensure that the moral law has objective content and is "designed expressly to be used at step (4) of the CI-procedure" (Rawls 2000, p. 222). This conception's dependence on the right is not obvious, since the specification of true human needs is independent of the right: what belongs to our true human needs is not

16 John Rawls, 2000, *Lectures on the History of Moral Philosophy*, ed. by Barbara Herman, Cambridge, Mass.: Harvard University Press. pp. 217–34.

17 Rawls's justification for ascribing this conception of the good to Kant is in Rawls 2000, 173 f., 221 f., 232–34.

determined (or restricted) by theories of the right but by theories of human nature. Yet the fact that true human needs are to be included into the theory of the good is a requirement of the CI-procedure and makes this conception of the good dependent on the theory of the right. The third conception of the good is the first which is without doubt not independent of the right: which ends are permissible and which impermissible cannot be determined without a conception of the right.[18] In Kant's theory the formula of humanity as end in itself would exclude as impermissible any conception of the good which involves using people merely as a means. The fourth, fifth and sixth conception of the good evidently presuppose a conception of the right. Hence, Kant does not define the good independently of the right and therefore accords the right priority over the good. Since there is in this respect no relevant difference between Kant's theory of the good in his earlier works and in the *Doctrine of Virtue* the priority of the right is true of the former and of the latter. Consequently, Wood's verdict that in the *Doctrine of Virtue* the good is prior to the right is incompatible with Rawls's definition of the priority relation between the right and the good. But Wood evidently employs a different conception of the priority of the right/good than Rawls. According to this conception, which is at least as widespread as Rawls' conception, the good has priority over the right if the right is somehow derived from the good. The right has priority if the principles of the right are justified without reference to or without presupposing principles of the good. It is commonly held that in this sense Kant and deontological theories give priority to the right whereas consequentialist and teleological theories give priority to the good. Recently, some Kant scholars have challenged this deontological reading of Kant in favour of a teleological reading which rejects the priority of the right. As an example of the traditional reading I will now discuss how Schneewind[19] contrasts the priority of the right in Kant with the priority of the good.

18 Kant mentions permissible ends in the *Doctrine of Virtue* [*MM* 6: 388 and 450].

19 J. B. Schneewind, 1992, "Autonomy, Obligation, and Virtue: An Overview of Kant's Moral Philosophy", in: Paul Guyer (ed.), *The Cambridge Companion to Kant*, Cambridge: Cambridge University Press, pp. 309–41.

II.

The two opposing views can be compared in the following table:[20]

Priority of the Good	Kant: Priority of the Right
(1) The goodness of states of affairs that can be brought about by human action is basic.	(1*) The concept of the good agent is basic.
(2) Being happy or having fully developed talents is intrinsically good.	
(3) A good agent is one who habitually and deliberatively does right acts.	(3*) A good agent is one whose will is wholly determined a priori by the moral law.
(4) A right act can be defined as one that brings about good states of affairs (to the greatest extent possible).	(4*) It is necessarily true that whatever acts a good agent does are right acts.
(5) We must know what is good before we can know what acts are right.	(5*) We do not discover what is right by first finding out what is good.
(6) Right acts have only instrumental value.	
(7) The goodness of a state of affairs occurs independently of the will of any finite moral agent (since that kind of state of affairs is intrinsically good because of the very nature of that kind of state of affairs).	
(8) If an agent must will to pursue good states of affairs, she is not self-legislating.	
	(9) Whatever states of affairs a good agent brings about are good states of affairs.
	(10) Only outcomes of right acts can count as good states of affairs.
	(11) We cannot determine what states of affairs are good without first knowing what is right.
	(12) In order to know what is right all we need to know is what the perfectly good agent would do.

20 The claims are taken literally or almost literally from Schneewind, 1992, p. 316 f.

To get a precise idea of each opposing view, we have to examine this contrast more closely. First, look at claims (4*) and (9) of the Kantian view. Taken together they yield:

(9') Whatever states of affairs a right act brings about are good states of affairs, i.e.,

Every outcome of a right act is a good state of affairs.

Since (10) is the converse of (9') and can be rephrased as

Every good state of affairs is the outcome of a right act.

we get:

(G) A state of affairs is a good state of affairs if and only if it is the outcome of a right act.

(G) can be taken as a definition of a good state of affairs. It presupposes that a right action can be defined independently of the notion of a good state of affairs, or else the definitions of a good state of affairs and right action would be circular. According to Schneewind's interpretation, then, Kant's theory has the following formal structure regarding the relation between the right and the good:

– Right action: Action with the property *P* (where *P* can be specified independently of the notion of a good state of affairs, e.g., being allowed by the Categorical Imperative).
– Good state of affairs: Outcome of a right action.

The opposing view reverses this structure and begins with a definition of a good state of affairs and then defines a right action in terms of its relation to a good state of affairs:

– Good state of affairs: State of affairs with the property *Q* (where *Q* can be specified independently of the notion of a right action, e.g., a state of affairs with the maximum amount of pleasure).
– Right action: Action whose outcome is a good state of affairs.

Instead of (G) this view holds:

(R) An action is right if and only if it brings about a good state of affairs.

Since this second view is the consequentialist view, Schneewind seems to describe the usual contrast between Kant and consequentialism.

	Consequentialism: Priority of the good	Kant: Priority of the Right
Right action	Action whose outcome is a good state of affairs	Action with the property *P* (identifiable independently of the notion of a good state of affairs)
Good state of affairs	State of affairs with the property *Q* (identifiable independently of the notion of a right action)	Outcome of a right action

Interestingly (as can be seen by looking at (4) and (9')), both views hold that a right action results in a good state of affairs, but each view attaches a different meaning to this:

– On the consequentialist view a right action results in a good state of affairs because that is how a right action is defined: an action is right only if its outcome is a good state of affairs.
– On Kant's view a right action results in a good state of affairs not because that is how a right action is defined, but because that is how a good state of affairs is defined: a state of affairs is good only if it is the outcome of a right action. Apart from right actions there are no good states of affairs.

While on the consequentialist view a good state of affairs need not be the outcome of a right action[21], on the Kantian view a good state of affairs can obtain *only* as the outcome of a right action. Taken literally, the Kantian view seems to be plainly wrong. It is a good state of affairs – obtaining independent of any action – that a country has a rich supply of raw materials or enjoys a moderate climate free from natural disasters. The outcome of an earthquake, on the other hand, is a very bad state of affairs. This way of speaking of good and bad states of affairs obtaining independently of any actions is natural and unobjectionable. A theory that cannot make sense of the notion of a good or bad state of affairs obtaining independently of any action is deficient.

A consequence of Kant's view, as it is interpreted by Schneewind, is that the consequences of actions can play no role in moral deliberations. Of course, in the past Kant was often credited with a complete disregard

21 The consequentialist definition of a right action as an action that brings about a good state of affairs entails that a good state of affairs cannot be the result of a wrong action. It is either the result of a right action or of no action at all.

for consequences, but today's Kantians are eager to show that this is a misunderstanding of Kant's ethics. Yet, if consequences are relevant in moral deliberation, they must be relevant as *good* or *bad* consequences, i.e., as *good* or *bad* states of affairs. But if one cannot tell good from bad states of affairs independently of right actions, consequences cannot be relevant in determining the right action. This entails that it does not make sense to hold that a prima facie wrong action may be right because it will prevent a disastrous consequence. Since the notion of a disastrous consequence, i.e., a very bad state of affairs, which can be specified prior to right actions is not available to Kant, it is conceptually impossible for him to say that the goodness or badness of a state of affairs can in some way determine whether an action which brings about this state of affairs is right or wrong. One cannot argue, e.g., that lying is wrong under ordinary circumstances, but right in this special circumstance where it is the only means to prevent that someone is being murdered. Consequently, Schneewind's interpretation of Kant entails a moral absolutism, according to which an action is right under every circumstance, no matter what the consequences. A further consequence of Schneewind's interpretation is that the so-called paradox of deontology is not applicable to Kant's theory. The paradox arises from "the rather simple thought that it can never be right to prefer a worse state of affairs to a better":

> we ask if it is not paradoxical that it should ever be morally objectionable to act in such a way as to minimize morally objectionable acts of just the same type. If it is a bad state of affairs in which one of these actions is done it will presumably be a worse state of affairs in which several are. And must it not be irrational to prefer the worse to the better state of affairs?[22]

This consequentialist objection against deontology has dominated the debate between deontology and consequentialism in recent years. Deontologists must address this objection, but on Schneewind's interpretation of Kant the objection is not even formulable because there is no notion of a good or bad consequence or state of affairs independent of right action. A real and deep conflict between deontology and consequentialism would thus be explained away.

Schneewind presents his account of the two opposing views as if these were the only possibilities. But that is obviously not the case. The Kantian and the consequentialist view do not exhaust the possibilities to relate the right to the good. They rather present two extreme positions compared

22 Philippa Foot, 1985, "Utilitarianism and the Virtues", in: Foot, *Moral Dilemmas*, Oxford: Clarendon Press 2002, pp. 59–77, at p. 62.

with the following two more moderate views, both of which acknowledge the notion of a good state of affairs specifiable independently of right actions:

Deontology 1:

- Right action: Action with the property *P* (identifiable independently of the notion of a good state of affairs)
- Good state of affairs: State of affairs with the property *Q* (identifiable independently of the notion of a right action).

This view would hold, in agreement with the consequentialist view, that good states of affairs do not depend on right actions but would disagree with both the Kantian and the consequentialist view in allowing that

- (contra Kant) not every right action results in a good state of affairs: A right action may bring about a bad state of affairs,
- (contra consequentialism) not every action that brings about a good state of affairs is a right action: A good state of affairs might be the result of a wrong action.

If right actions and good states of affairs are defined independently of each other, these relationships between right actions and good states of affairs are only to be expected. It would be a peculiar coincidence if every right action would bring about a good state of affairs or if no wrong action could bring about a good state of affairs. Some might object to the view that a wrong action could bring about a good state of affairs and would want to hold that, e.g., although the well-being of my family is *ceteris paribus* a good state of affairs, it is *not* a good state of affairs if it was made possible by first committing a wrong action, like robbing a bank or murdering someone to inherit the money. Those inclined to this view might want to modify *Deontology 1* to *Deontology 2:*

- Right action: Action with the property *P* (identifiable independent of the notion of a good state of affairs)
- Good state of affairs: State of affairs with the property *Q* (identifiable independent of the notion of a right action) *and* not the result of a wrong action.

This view would, again in agreement with the consequentialist view, hold that there might be good states of affairs independently of right actions – as long as they are not the result of wrong actions. And it would, like *Deontology 1*, allow that a right action may bring about a bad state of

affairs. But, unlike *Deontology 1*, it would claim that no wrong action can result in a good state of affairs.
The four views can be compared in the following table:

	Consequentialism	*Deontology 1* (Moderate Deontology)	*Deontology 2* (Moderate Deontology)	Kant, Absolutism (Extreme Deontology)
Right action	Action whose outcome is a good state of affairs	Action with property P^{23}	Action with property P^{23}	Action with property P^{23}
Good state of affairs	State of affairs with property Q^{24}	State of affairs with property Q^{24}	State of affairs with property Q^{24} *and* not the result of a wrong action	Outcome of a right action
	Priority of the good			Priority of the right

We can now, based on Schneewind's account, give a definition of the priority of the right/good:
The right has priority over the good if

- the right is defined without any reference to or presupposing anything about the good, *and*
- the good is defined completely in terms of the right.

The good has priority over the right if

- the good is defined without any reference to or presupposing anything about the right, *and*
- the right is defined completely in terms of the good.[25]

23 *P* is identifiable independently of the notion of a good state of affairs.
24 *Q* is identifiable independently of the notion of a right action.
25 The second condition in each definition is necessary because the two first conditions alone do not exclude each other and can (as in *Deontology 1*) be fulfilled together. Without the second condition one would have to allow that the right has priority over the good and the good has priority over the right, which does not make much sense.

This definition of the priority of the right (good) might be interpreted as meaning that the right (good) is the *only*[26] *basic concept.* This interpretation actually does not accord with Schneewind's view, because according to him the concept of the good agent (rather than the concept of the right) is basic in Kant's theory (cf. claim (1*) above). But this view has its problems: if the *good* agent is the basic concept, it seems that in Kant's theory the *good*, rather than the *right* is the basic concept – just like in consequentialism, where the *good* state of affairs is the basic concept. But there is a crucial difference between Kant and consequentialism which precludes this conclusion: in consequentialism the good is basic because a good state of affairs is defined independently of the right and the right is then defined in terms of the good. Correspondingly, only if the good agent is defined independently of the right and the right then defined in terms of the good agent, would the good be the basic concept in Kant's theory. But that is not the case. According to Schneewind, we are supposed to think of "the good agent as one whose will is wholly determined *a priori*, and [to] think of the pattern of that determination as the moral law" (Schneewind 1992, p. 316). If we equate the moral law with the right, the good agent is not defined independently of the right and therefore the good is not the basic concept in Kant's theory. The good agent cannot be the basic concept either: if the good agent is the basic concept in Kant's theory in the same way as the good state of affairs is the basic concept in consequentialism, the right in Kant's theory must be defined in terms of the good agent (just as consequentialism defines the right in terms of the good state of affairs). But we just saw that the good agent is *not* defined independently of the right and therefore cannot be the basic concept. Rather, the right is the basic concept in terms of which the good agent is defined. Maybe Schneewind believes that the right is defined in terms of the good agent when he writes that "[i]n order to know what is right all we need to know is what the perfectly good agent would do" (Schneewind 1992, p. 317).[27] But how does the perfectly good agent herself know what is right? She must have a conception of the right that does not depend on the concept of the good

26 In theories without a priority relation, like *Deontology 1*, the right and the good are both basic concepts.

27 This may be taken as a harmless consequence of the claim that whatever acts a good agent does are right acts (Schneewind 1992, 316): if a perfectly good agent always does right acts, we know what is right when we know what such an agent would do.

agent. And indeed, Kant's perfectly good agent does have such a conception of the right, namely the Categorical Imperative. Thus, in spite of Schneewind explicitly saying that the concept of the good agent is basic in Kant's theory, this view must be rejected: the moral law and therefore the right is the basic concept in terms of which the good agent is defined.[28]

The above-mentioned definition entails that the two priority relations between the right and the good are not contradictories but subcontraries: it is not the case that either the right or the good must have priority. There are theories, like *Deontology 1* and *Deontology 2* in which neither the right nor the good has priority. Hence, one is not forced to choose between the priority of the good or the priority of the right. Rejection of the consequentialist view does not commit to the view ascribed to Kant. A further consequence of this definition is that the traditional characterization of deontological theories as theories where the right has priority over the good is untenable: as can be seen in the above table, in deontological theories either the right has priority over the good, or there is no priority relation between the right and the good. Furthermore, since most deontological theories, even, as we shall see, Kant's theory, belong to either of the two moderate versions of deontology, and possibly nobody holds the extreme view which Schneewind ascribes to Kant, Schneewind's definition robs the priority of the right any significance.

But what's more, Schneewind does not give a true account of Kant's theory because Kant does have a conception of the good which is independent of the outcome of a right action.[29] Like consequentialists, Kant can regard happiness as good. The only – important – difference is that for Kant it is not unconditionally good. It is good only if accompanied by a good will so that

> an impartial rational spectator can take no delight in seeing the uninterrupted prosperity of a being graced with no feature of a pure and good will, so that a good will seems to constitute the indispensable condition even of the worthiness to be happy [*G* 4: 393].

28 Actually, I cannot see why consequentialists and Kant should have a different concept of the good agent. They both share the same concept of a good agent, and the consequentialist can agree to Schneewind's claims (3*) and (4*) as well as Kant can agree to claim (3).

29 This conception corresponds to the first two conceptions of the good in Rawls's account of Kant's theory of the good.

Thus, the good will and right action (since a morally good person necessarily does right actions) seem to be a necessary condition for something to be good. But still, Kant would not count *everything* that is accompanied by a good will as good. He certainly would not say that the unhappiness of a virtuous person is good – on the contrary. Thus, Kant must have a conception of the good prior to and independently of the moral law – a conception of the good which allows him to distinguish between those states of affairs which can be good if accompanied by a good will and those which cannot be good, even if accompanied by a good will. Another way of saying this is that Kant must have a conception of the non-moral good, the good from a non-moral point of view, and a conception of the moral good, the good from the moral point of view. From this latter point of view a non-morally good state of affairs is good only if accompanied by a good will. If this is true, Kant's theory must be classified as *Deontology 2* and does not give priority to the right in Schneewind's sense.

III.

The reason for thinking it necessary to interpret the priority of the right in Schneewind's extreme sense and attributing it to Kant lies, I believe, in a very common misunderstanding of consequentialism.[30] Consequentialism need not be interpreted as having only one basic concept, namely the good. The definition of a right action as one that brings about a good (or rather the best) state of affairs does entail neither that the very *concept* of the right nor that the *content* of the right (i.e., this definition of a right action) must be derived from the good. Consequentialism can have two basic concepts, the right and the good, and like Sidgwick one may regard the definition of the right as the maximization of the good as a basic (autonomous) intuition of the right that is not derived from a prior concept of the good. This possibility is overlooked because Schneewind, along with many other philosophers, believes that if the right is defined as the maximization of the good it is defined instrumentally as means to bringing about good states of affairs such that right actions can only have instrumental value. As Freeman puts it:

30 With reference to Schneewind's claims in the table above, the misunderstanding is that consequentialists are committed to the claims (6), (7) and (8).

> […] teleological views (1) define the good independent of any moral concepts; and then (2) define the right purely in instrumental terms of principles of expedience, i.e., as what most effectively and probably realizes the greatest amount of good.[31]

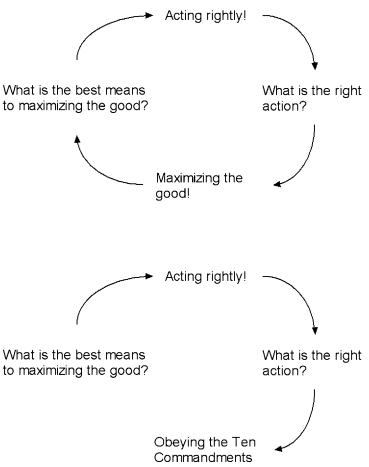
I will now argue that this view is mistaken, i.e., it is wrong to suppose that if the right is defined as that which maximizes the good, then it is defined as instrumental to the good. My general point will be that in consequentialism the right *consists* in the maximization of the good but is not the *means* to the maximization of the good.

If something is instrumental to some other thing it is a means to that thing. If, therefore, the right is defined as instrumental to the maximization of the good, the right must be a means to the maximization of the good. Further, if the right is defined as that which maximizes the good, it must be the *best* means to the maximization of the good. If an action were not the best means, there would be some other action, which is the best means and therefore produces more good and would thus be the right action, contrary to the assumption that the first action was the right action. The instrumental definition of the right, then, amounts to the claim: an action is right if and only if it is the best means to maximizing the good. But this definition is not equivalent to the consequentialist definition that an action is right if and only if it maximizes the good. Maximizing the good is not equivalent to being the best means to the maximization of the good because (i) an action that is the best means to maximize the good need not in fact maximize the good, and (ii) an action that in fact maximizes the good need not have been the best means to maximize the good. Hence, defining the right as the maximization of the good does not entail that the right is the best means to the maximization of the good.

The question whether something is the best means to the maximization of the good is an empiricial question which cannot be decided by philosophical arguments. One cannot *define* the best means to the maximization of the good. If the right were defined as, e.g., the Ten Commandments or Ross's list of prima facie duties or Kant's Categorical

31 Samuel Freeman, 1994, "Utilitarianism, Deontology and the Priority of the Right", *Philosophy and Public Affairs* 23, pp. 313–49, at p. 323 f. Similar views are held, e.g., by Darwall, 2003, p. 2 f.; Christine M. Korsgaard, 1998, "Good, Theories of the", in: Edward Craig (ed.), *Routledge Encyclopedia of Philosophy*, vol. 4, London: Routledge, pp. 130–35, at p.133; and Jens Timmermann, 2005, "Why Kant Could not Have Been a Utilitarian", in: *Utilitas* vol. 17, pp. 243–64, at p. 258.

Imperative, one can ask whether, as a matter of fact, obeying these principles is the best means to the maximization of the good. If, instead, the right is defined as the maximization of the good, the above question involves a circle: what is the best means to the maximization of the good? Acting rightly! What is the right action? Maximizing the good![32]



Consequently, the question whether the right is the best means to the maximization of the good makes sense only if it is *not defined* as the maximization of the good. Only if one has an independent definition of the right, one can sensibly ask whether the right, thus defined, is the best means to the maximization of the good. When asked what is the best means to the maximization of the good, consequentialists do not answer "Acting rightly!" Rather, they answer, e.g., that due to the diminishing marginal utility of some goods the best means to maximize the good is to distribute these goods equally. (Thus, equal distribution, not right action, has instrumental value as a means to maximize the good.)

If the right would not consist in the maximization of the good, but would be the means to the maximization of the good, we should do the right thing only if we want to maximize the good. If we do not will the end we do not need to will the means. Thus, the instrumental definition of the right as a means to the maximization of the good yields only the

32 Consider instead: what is the best means to the maximization of the good? Acting Rightly! How can I act rightly? Act in accordance with the Categorical Imperative! (Or: Obey the Ten Commandments!)

hypothetical imperative "If you want to maximize the good, act rightly!" and would reduce morality to hypothetical imperatives.

Once it is appreciated that consequentialism is not tied to the view that the right is a derivative concept having only instrumental value one sees that the difference between consequentialism and Kant or deontology is not a difference between the priority of the good and the priority of the right. Both theories employ the same concepts of the good and the right. There is no need to reject consequentialism in order to have an autonomous conception of the right. If the right is not instrumental to the promotion of the good, then consequentialism, too, has an autonomous conception of the right and regards the maximization of the good as a categorical imperative. With regard to Kant's arguments against the priority of the good in the *Critique of Practical Reason*, we can draw the following conclusion: if the right in consequentialism is a categorical imperative and not instrumental to the promotion of the good, the right, rather than the good is the determining ground of the will. And that is exactly how Kant understood the priority of the right: the determination of the will by the moral law. We are now in a position to see the priority relation between the right and the good in a different light. In many discussions of the priority relation it is assumed that it describes a fundamental difference between consequentialism and deontology. But in fact, in both consequentialism and deontology the right is prior to the good in Kant's sense because both maintain that the right is a categorical imperative which we are to obey regardless of our inclinations and our personal good. The difference between the priority of the right and the priority of the good is not the difference between deontology and consequentialism but the difference between what Charles Larmore[33], following Sidgwick, has called the imperative view (of modern ethics) and the attractive view (of ancient ethics), or, as one might say, the ethics of duty and the ethics of virtue. Whereas in the ethics of virtue the good life consists in the virtuous life, the priority of the right opens a gap between an agent's personal good and the demands of morality. This is true of Kantian deontology as well as of consequentialism, where the good to be promoted is not the agent's own good.

As regards his own conception of the priority of the right, Kant did not change his mind between the *Groundwork* and the *Doctrine of Virtue.*

33 Charles Larmore, 1996, *The Morals of Modernity*, Cambridge: Cambridge University Press, ch. 1.

Thus, although Wood's claim that Kant "grounds the duty to act in certain ways exclusively on the promotion of ends" may be true, that still leaves Kant on the side of the priority of the right as long as he regards the duty to act in certain ways as a categorical imperative.

Being Virtuous and the Virtues: Two Aspects of Kant's *Doctrine of Virtue*

Philip Stratton Lake

An account of virtue should include an account of what it is for a person to be virtuous and an account of the various specific virtues. In addition to this it should show how these different aspects of virtue are related to each other. This additional task may be very straightforward. One might offer an account of each of the virtues and maintain that to be virtuous is just to have all, or most, of these virtues. Alternatively, one might define each virtue as a sensitivity to a specific range of moral considerations or reasons, and define a virtuous person as one who has a sensitivity to all or most of these considerations. It would be a deficiency in a theory of virtue if it failed to establish an intimate link between what I shall henceforth refer to as the *two aspects of virtue* – that is between the specific virtues and being virtuous as such. It would be a serious flaw in such a theory if it meant that an agent's virtue *could* not be realized in the exercise of one of her specific virtues.

In this paper I want to outline Kant's account of the two aspects of virtue in his *Doctrine of Virtue.* Having outlined these two aspects of his doctrine of virtue I shall consider how these different aspects are related. I shall argue that on Kant's view virtue (the state of being virtuous) cannot be realized in the exercise of the specific virtues, and that this is a major deficiency in his doctrine of virtue. I shall then consider how his account might be revised so that these two aspects of virtue may be appropriately related to each other.

I. Kant's Account of Virtue

Kant's account of what it is to be a virtuous person is based on his account of morally worthy acts in the *Groundwork.* In the *Groundwork* Kant notoriously claimed that only actions done from duty have moral worth. Acting from duty, for him, is not acting from any inclination, and so is not acting from an inclination to do what is right. It is not, therefore,

acting in pursuit of some desired end. Since acting from duty is not acting in pursuit of some end, the value of actions done from this motive do not depend on success in the attainment of any end. So our willing would retain its moral worth even if, through some misfortune, we failed to achieve some intended end, or even if we failed to do anything at all *G* 62 [4: 394].[1]

This claim about moral goodness is sometimes thought to imply that Kant believed that morally right (obligatory) acts are independent of any end, and thus are independent of any success in the attainment of some end.[2] But this interpretation conflates a distinction Kant clearly makes between morally good and morally right acts – that is, between moral worth and obligation. A morally right act is one that accords with duty. A morally good act is one that is done from duty. Since one can act in accordance with duty without acting from duty[3], one's act can be morally right yet lack moral worth. So moral goodness and moral rightness are distinct notions in Kant. Consequently, the fact that he thinks that the moral worth of an act is independent of success in the attainment of one's end does not imply that he thought that the moral rightness of an act is independent of such success.

In fact there is good reason to think that Kant thought that moral rightness does depend on such success. Kant thinks that we have a duty to keep our promises. What it is we ought to do in such cases is determined by what we have promised to do. Typically we do not promise to *try* to do certain acts, but promise to actually do them. In such cases we have done what we should only if we succeed in doing those actions.[4] We have not done what we should if we merely will, but fail, to do them.[5] Although

---

1 Immanuel Kant, *Groundwork of the Metaphysic of Morals*, trans. by H. J. Paton, 1964, New York: Harper and Row (hereafter abreviated as *G*).

2 See, for example, H. J. Paton, 1947, *The Categorical Imperative: A Study in Kant's Moral Philosophy*, London: Hutchinson's University Library, p.58.

3 Kant's examples of the shop-keeper who is honest from self-interest and of the "friend of man" who is beneficent from inclination clearly illustrate this.

4 This is subject to certain conditions. The promisee may cancel this obligation by letting us off the promise. There will also be other implicit conditions that need not figure in the content of the promise.

5 Under the influence of Prichard, W. D. Ross came to think that all we are bound to do is try, or as he put it, "set ourselves", to do certain acts. We cannot, he thought, be required actually to do them (W. D. Ross, 1939, *The Foundations of Ethics*, Oxford: Clarendon Press, pp. 148 ff). But in *The Right and the Good* (W. D. Ross, *The Right and the Good*, 2002, Oxford: Clarendon Press, pp. 42 ff) he

we would judge someone less harshly if they tried but failed to do what they promised, in so far as they fail to do what they have promised they have still failed to do what they should, and so would owe the promisee an apology. Unlike morally good acts, then, morally right acts sometimes involve success in the pursuit of one's aims.

Now one might act from duty even though one is not a virtuous person. Nonetheless, Kant's account of moral worth in the *Groundwork* suggests an account of virtue as a general *disposition* to act from duty. This suggestion is confirmed in the *Metaphysics of Morals.* Here Kant distinguishes between formal and material ethical duties, and maintains that there is only one formal, but several material ethical duties. He writes:

> Man has a duty to carry the cultivation of his *will* up to the purest virtuous disposition, in which the *law* becomes also the incentive to his actions that conform with duty and he obeys the law from duty. This disposition is inner morally practical perfection (*MM* 191–2 [6: 387]).[6]

And later on he writes that:

> Man's greatest moral perfection is to do his duty *from duty* (for the law to be not only the rule but also the incentive of his actions) (*MM* 196 [6: 392]).

To attain the greatest moral perfection is to acquire the purest virtuous disposition, and this is a general disposition to act from duty. Later on Kant repeats this account of virtue (of being virtuous) when he writes that "virtue is not merely a self-constraint… but also a self-constraint in accordance with a principle of inner freedom, and so through the representation of one's duty in accordance with its formal law" (*MM* 197 [6: 394]).

It seems then that, in the *Metaphysics of Morals*, Kant develops an account of virtue from his earlier account of morally worthy actions. On this reading, morally worthy actions are primary, and the notion of virtue is defined with reference to this primary notion.

However, this reading is complicated by the fact that Kant also defines the notion of a virtuous action with reference to the agent's virtue – that is, as an action that realizes the agent's virtue as the disposition to act from duty.

---

provided good reasons for rejecting what was to become his later view, and never tells us why he no longer found these objections compelling.

6 Immanuel Kant, *Metaphysics of Morals*, trans. by M. Gregor, 1991, Cambridge: Cambridge University Press (hereafter abbreviated as *MM*).

> Since the moral capacity to constrain oneself [through a law] can be called virtue, action springing from such a disposition (respect for law) can be called virtuous (ethical) action… (*MM* 198 [6: 394]).

This seems to reverse the priority of the *Groundwork* where Kant seems to think that the moral worth of particular acts is primary and independent, and that a good will is understood as a disposition to do such acts. But we may make Kant's later view compatible with his earlier one if we distinguish between a virtuous act and a morally worthy act. An action is morally worthy if and only if it is done from duty, according to Kant. But one can act from duty even if one is not virtuous (one might be trying to be virtuous, though one has not yet achieved this goal). If one is not virtuous, yet acts from duty, one's action cannot be the realization, or expression, of one's virtue, and so cannot be a virtuous act as Kant describes this in the *Metaphysics of Morals.* But since acting from duty is sufficient for moral worth, such an action would have moral worth even though it was not a virtuous action. Acting from duty is, then, a necessary, but not a sufficient condition of virtuous acts. Virtuous actions must not only be done from duty, but must also realize, or express, a general disposition to fulfil all of one's obligations from duty.

There is a trivial sense in which every action done from duty will be the realization of a disposition to do such acts. For if I do an act of a certain sort in a certain set of circumstances, then there must be a sense in which I am disposed to do such acts under those conditions. If this is right, then whenever I do some act from duty it will be true that I am disposed to do that act from duty, and the difference between a morally worthy and a virtuous act will be lost.

But this problem is generated by a failure to distinguish a very specific disposition from a much more general one. I may do some specific act from duty, and so may be disposed (in some weak sense) to do such acts in such circumstances. But the sort of disposition Kant has in mind as constituting virtue is a much more general disposition to act from duty. It is a disposition to do any action that is recognized as obligatory just because it is obligatory. I may do some specific act from duty but lack this general disposition. In such cases my act will be morally worthy but not virtuous.

Although there is only one virtuous disposition – the (general) disposition to act from duty – there are several virtues corresponding to the several duties of virtue.

> Those duties that have to do not so much with a certain end (matter, object of choice) as merely with *what is formal* in the moral determination of the will (e.g., that an action in conformity with duty must also be done *from duty*) are not duties of virtue. Only *an end that is also a duty* can be called a *duty of virtue.* For this reason there are several duties of virtue (and also various virtues), whereas for the first kind of duty only one (virtuous disposition) is thought, which however holds for all actions (*MM* 188 [6: 383]).

> Like anything *formal*, virtue as the will's conformity with every duty, based on a firm disposition, is merely one and the same. But with respect to the *end* of actions that is also a duty, that is, what one *ought* to make one's *end* (what is material), there can be several virtues; and since obligation to the maxim of such an end is called a duty of virtue, there are many duties of virtue (*MM* 198 [6: 395]).

Sometimes Kant calls the single notion of virtue an *obligation of virtue*, and contrasts this with the plurality of *duties of virtue.* He thus writes that

> there is only *one* obligation of virtue, whereas there are *many* duties of virtue; for there are indeed many objects that it is also our duty to have as ends, but there is only one virtuous disposition, the subjective determining ground to fulfil one's duty… (*MM* 210 [6: 410]).

What we get from these passages is a distinction between being virtuous in general – that is, being a virtuous person – and the various virtues. Virtue as such is understood as a disposition to act from duty, i.e., from respect for law. This single disposition is to be contrasted to the plurality of virtues, of beneficence, appreciativeness, compassion, etc., which are determined by compliance with the various duties of virtue.

This division is complicated somewhat by the fact that Kant thinks that one of the duties of virtue requires us to perfect ourselves. Since our moral perfection consists in our being disposed to act from duty, the one obligation of virtue seems also to be a duty of virtue. So Kant seems to end up identifying the one state of being virtuous with one of the several virtues with which it is contrasted. If this is right, then the two aspects of virtue will end up being one.

To avoid this we have to distinguish between striving to make oneself virtuous and the state of virtue one is striving to attain. What we aim at when we strive to make ourselves virtuous is the single state of being virtuous: which according to Kant is the general disposition to treat duty itself as a sufficient ground of action. This disposition is merely formal: it is not a disposition to pursue some end, but to make the law itself the sufficient ground of all moral action. To *strive* to make ourselves virtuous is, however, to make this (formal) disposition our end. It is not to be

virtuous, but is to have one of the several virtues Kant mentions. Virtue then is the disposition to make duty the sole incentive of right acts. The struggle to acquire this disposition is a specific virtue, a virtue we possess in so far as we make our perfection our end.

II. Imperfect Duties and the Virtues

Kant bases his account of the various virtues on conformity with duties of virtue. Duties of virtue are imperfect duties (*MM* 194 [6: 390]). The notion of an imperfect duty can be traced back to Grotius and Pufendorf. For Grotius the distinction between perfect and imperfect duties was primarily a distinction between rights. For him, perfect rights are those that licence their possessor to use force in their pursuit, where necessary, whereas imperfect rights do not carry this entitlement.[7] Pufendorf develops this distinction in terms of duties. Like Grotius, he maintained that compliance with perfect duties may be coerced, whereas compliance with imperfect duties may not. He adds to this that perfect duties are specific, requiring, for instance, that we pay a specified individual a precise amount, whereas imperfect duties, such as the duty of beneficence, do not specify whom we should benefit or how much we should give. He also claims that perfect duties are essential to the existence of society whereas imperfect ones are not.

Kant first introduces his distinction between perfect and imperfect duties very briefly in the *Groundwork*. Imperfect duties are there distinguished from perfect duties in two ways. They are distinguished by what we might call their "deontic valence" and their "deontic force".

---

7 "The second way is, when future intentions are expressed by outward acts and signs sufficient to indicate a resolution of abiding by present assurances. And these kind of promises may be called imperfect obligations, but conveying to the person to whom they are given no RIGHT to exact them. For it happens in many cases that we may be under an obligation of duty, to the performance of which another has no right to compel us. For in this respect the duty of fidelity to promises, is like the duties of compassion and gratitude. In such kinds of promises therefore the person to whom they are made, has no right, by the law of nature to possess himself of the effects of the promiser, as his own, nor to COMPEL him to the performance of his promise." (*On the Law of War and Peace,* Book Two, Ch.II. iii). For a good, brief overview of Grotius's and Pufendorf's views see Jerome Schneewind, 1990, "The Misfortunes of Virtue", in: *Ethics*, vol. 101, pp. 42–63.

The deontic valence of a duty is determined by whether it prescribes or proscribes some act, whether it states that some act is required or wrong. Perfect duties proscribe certain actions – that is, they state which acts we must not do. Examples of perfect duties are the prohibition on lying and suicide. But knowing what it is we must not do does not allow us to infer what we must do. For not all acts that are not morally forbidden are morally required: some are neither forbidden nor required. Imperfect duties tell us which of the actions that are not wrong are required. In the *Groundwork* Kant gives us two examples of imperfect duties: the first is the requirement to develop our talents, and the second is to make others' ends our own.

The second feature of imperfect duties in the *Groundwork* seems to distinguish them by their deontic force (but as we shall see shortly, this impression is mistaken). In a footnote Kant states that perfect duties allow of no exceptions in the interest of inclination.[8] This suggests that he thinks that imperfect duties have less force, i.e., are less strict in the sense that they do allow of exceptions in the interest of inclination. But how are we to make sense of the idea of an obligation that we may ignore if we feel like it? Such an obligation seems to have no force at all.

We can make no sense of this second feature of imperfect duties if we think of them as requiring that we do certain actions. For then Kant would be claiming that we must do such actions, though we don't have to do them if we don't want to. Pufendorf's claim that imperfect duties lack specificity does not help here. Even if imperfect duties do not tell us how much we should do to help, or whom we should help, they will still require that we help; and it is difficult to see what this amounts to if we can excuse ourselves from helping on grounds of inclination. If, then, we are to make sense of Kant's claim in the *Groundwork* we must assume that he had in mind something like the view laid out in the *Metaphysics of Moral*, namely, that imperfect duties do not require us to do certain acts but require us to adopt certain ends.

Adopting an end, for Kant, means adopting the relevant maxim of ends. A maxim is a subjective principle of action.[9] To adopt a maxim of ends is, then, to make it our principle that we will make certain ends our own. But to have some end does not mean that we take every opportunity to pursue it. I could have the goal of keeping fit without taking every opportunity to exercise. I would still have this end if, for instance, on a

8 *G* 89n [4: 421n].
9 *G* 69n [4: 400n].

rainy afternoon, I have the time to go for a run but don't, because I don't fancy getting wet. Doing nothing in pursuit of this end would clearly suggest that I do not really have it, and if I sincerely claim that I do, then I would probably be suffering from self-deception. But there is plenty of space between this extreme and the other extreme of taking every opportunity to exercise. This space allows for different degrees of commitment to this end. If I take very few opportunities to exercise, then I might have this end, but have only a weak commitment to it. If, on the other hand, I take every opportunity to pursue it I will have a very strong commitment to this end.

It is this feature of possessing an end that allows for latitude. If I am required to make others' ends my own then I have done this even if I do the minimum that is necessary to count as having it or if I take every opportunity to help others.

> …if the law can prescribe only the maxim of actions, not actions themselves, this is a sign that it leaves a latitude (latitude) for free choice in following (complying with) the law, that is, that the law cannot specify precisely in what way one is to act and how much one is to do by the action for an end that is also a duty (*MM* 194 [6: 390]).

It turns out then that what, in the *Groundwork*, looked like a difference in deontic *force* is in fact a difference of deontic *content.* Perfect duties prohibit and imperfect duties require quite different things. Perfect duties forbid the performance of certain actions whereas imperfect duties require the adoption of certain maxims (maxims of ends). This distinction does not commit Kant to the view that these different duties have a different deontic force, and once we are clear that these different types of duty have different objects (content) we have no reason to ascribe this view to Kant. Imperfect duties require us to adopt certain maxims with the same force that perfect duties prohibit certain actions.

We have seen that imperfect duties allow for a certain latitude because they require the adoption of certain ends, rather than the performance of certain actions. In the *Groundwork* Kant makes it sound as though this latitude enables us not to help others in certain circumstances if we are not inclined to do so. In the *Metaphysics of Moral*, however, latitude is not given in the interest of inclination, but in the interest of pursuing some other imperfect duty.

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

> which in fact the field for the practice of virtue is widened (*MM* 194 [6: 390]).

His later view is, therefore, more demanding than his earlier view, and in some ways more plausible. We could imagine someone who (to borrow Kant's own example in the *Groundwork*) had adopted a maxim of beneficence, but was disinclined to help someone in need because he was temporarily absorbed with his own problems. It seems wrong to say that this person had done all that morality requires of him (adopted a beneficent maxim) even though he failed to help when he could easily have done so. Kant's view in the *Metaphysics of Morals* would be that this person had not done everything that morality required of him, whereas the view he suggests in the *Groundwork* would seem to allow that he had. (Though we will see that his later view is not without its difficulties.)

So far we have characterised Kant's distinction between perfect and imperfect duties with reference to their valence and content. Perfect duties proscribe certain actions whereas imperfect duties prescribe certain maxims. A third feature of imperfect duties relates to their coercibility.

Grotius held that perfect rights entitle their possessor to use force in their pursuit where necessary, whereas imperfect rights carry no such entitlement. Kant makes a similar point in relation to perfect and imperfect duties, though he differs in certain respects. First, he thought that compliance with imperfect duties not only *may* not be compelled, but that it *could* not be compelled. We could compel people to act as if they have adopted certain ends, but could not compel them to adopt those ends (*MM* 187 [6: 381]). For Grotius the difference between perfect and imperfect rights rests on a normative claim. With perfect rights the agent is *entitled* to coerce compliance, whereas with imperfect rights he is not. Kant's distinction does not rest on this, or any other normative claim, but on the non-normative claim that we are *able* to coerce compliance with perfect duties, whereas we are not able to do so with imperfect duties.

Second, Kant did not think that all perfect duties could be compelled. For instance, perfect duties to oneself cannot be coerced. Furthermore, saving one's own life by violence is contrary to a perfect duty to others, but could not be deterred by the threat of punishment, for the agent's certain death will always give him a stronger incentive than the threat of punishment (*MM* 60 [6: 235–6]).

A fourth distinction between perfect and imperfect duties relates to merit. Compliance with perfect duties that can be compelled (duties of

Right) lacks merit. Kant offers two reasons for this claim. First, in so far as I comply with such duties I do merely what is owed (*MM* 53 [6: 227]). Failure to do what is owed is demerit, but doing what is owed is merely doing what others are entitled to compel from me, and so is not in itself praiseworthy. If I return what I have borrowed from you I have complied with a perfect duty. My action is, therefore, right but not meritorious, for I have merely done what is owed to you and what you are entitled to compel me to do. This suggests that when I comply with imperfect duties I go beyond what I owe to others, and thus go beyond what others are entitled to compel me to do.

This feature of imperfect duties is much closer to Grotius' normative claim about coercibility. What makes compliance with imperfect duties meritorious is not that others *cannot* compel compliance, but that they are not *entitled* to compel compliance. They are not entitled to compel compliance because compliance is not owed. What determines what we owe to others is left unclear, and I do not wish to pursue this here.[10] All I wish to highlight is that one of the reasons Kant thinks that compliance with imperfect duties is meritorious is because we go beyond what we owe to others, and thus beyond what we may legitimately be compelled to do.

The second reason why compliance with imperfect (but not perfect) duties is meritorious is that in complying I do more in the way of duty than I *can* be coerced to do (*MM* 53 [6: 228]). Furthermore, compliance is meritorious to the degree that I do more than I can be coerced to do.

> The greater the natural obstacles (of sensibility) and the less the moral obstacles (of duty), so much the more merit is to be accounted for a *good deed*, as when, for example, at considerable self-sacrifice I rescue a complete stranger from great distress (*MM* 53 [6: 228]).

Such actions are not supererogatory in the sense that they are beyond duty. They are, however, beyond duties of Right, and so bear some resemblance to supererogatory acts.

This is a distinct ground of merit from the first. Here, Kant is focusing on the third feature of imperfect duties, mentioned above, that they cannot be backed by coercive force. This is quite different from the other ground of merit already mentioned, that we are not entitled to back compliance with coercive force. This non-normative consideration is less

10 One suggestion would be that we have a right not to have their autonomy undermined, either by force, or deception, but we do not have a right against others that they promote our autonomously chosen ends.

plausible as a ground of merit, for it is not clear why the fact that we can be coerced to comply with perfect duties means that compliance lacks merit? First, the fact that a perfect duty can be backed by coercive force does not imply that it is so backed. In cases where such force is absent it is unclear to me why we should claim that compliance lacks merit. Second, even if a perfect duty were backed by coercive force there is no reason to suppose that we would always comply from fear of punishment. In cases where we do not comply from fear, I see no reason (distinct from the first reason) why we should deny merit. So of the two grounds of merit Kant offers, doing more than what is owed and doing more than can be coerced, it is only the first that may be regarded as plausible.

We may summarise the various aspects of Kant's account of virtue as follows: The various virtues are grounded in compliance with duties of virtue. These duties are imperfect, that is, they are duties that prescribe certain maxims and so allow latitude in complying with them; compliance cannot be coerced, and compliance is meritorious because it goes beyond what is owed. To be virtuous, or to have virtue is not, for Kant, merely to comply with all of these duties, but is to be disposed to comply with all duties from duty. Finally a virtuous act is one that is the realization of this disposition. It is to do some act solely from duty. All instances of acting form duty have moral worth, but not all such actions are virtuous. For one might act from duty even though one lacks virtue. Since such actions cannot be the realization of virtue, they cannot be virtuous acts.

III. The Virtues and Motivation

It is generally accepted that motivation plays an essential role in any account of the virtues. To have a some specific virtue is not just to act in certain ways, but to do so for the right reason. We have seen that acting from some specific motive (duty) exhausts Kant account of what it is to be a virtuous person. The question we need to address now is whether his account of the specific virtues includes a motivational element, and if so what the details of this are.

For Kant, to have some specific virtue is to act in accordance with the relevant duty of virtue. So, for example, to be beneficent is to act in accordance with the imperfect duty of beneficence. This duty requires us to make others' ends our own. If Kant's account of the virtue of beneficence involves a motivational element we may assume that this

motive will be the motive of duty. So the duty of beneficence would require not only that we make others' ends our own, but that we do this from duty. On this account of the duty of beneficence, we fail to comply with this duty if we make others' ends our own from some other motive, say, from self-interest (indirect egoism), or from inclination. For then we would have done only part of what we ought. Consequently, we could not be said to have the virtue of beneficence, or if we could, it could only be a defective, or incomplete sense in which we do.

This view of duties of virtue is supported by Kant's claim that all ethical duties require not only the adoption of an end, but also that the law prescribing this end be the determining ground of that act (*MM* 46 [6: 218–9], *MM* 42 [6: 214]), and many commentators read Kant in this way.[11] But Mark Timmons has recently argued that these passages do not support the view that the motive of duty is part of the content of duties of virtue. He rejects the view that, for Kant, duties of virtue require us to adopt certain ends and to adopt them *from duty*.[12] As he understands Kant, the motive of duty is not part of the content of duties of virtue, but is a *precondition* of those duties (Timmons 2002, 270).

His reason for this interpretation is premised on the assumption that Kant is an ethical internalist. Ethical internalists maintain that we can only be obligated to do some act if we have a sufficient motive for complying (*ibid*).

> The basic idea is that in order for an individual to be morally required to perform some action, it must be true of the agent that (1) there is an all-things-considered good normative reason for her to perform the action in question, which (2) is motivationally available to her (*ibid*).

In the case of duties of Right an incentive is provided by the threat of punishment. But because the adoption of ends cannot be coerced, this incentive is not available for ethical duties. Some other incentive is required. This incentive cannot be provided by some desire, as this would make these duties depend upon contingent desires. So the incentive must be the law (duty) itself. The motive of duty is not part of what the law

11 Onora O'Neill, 1975, *Acting on Principle: An Essay on Kantian Ethics*, New York: Columbia University Press, chs. 4–6; and Barbara Herman, 1993, *The Practice of Moral Judgment*, Cambridge Massachusetts: Harvard University Press, p. 15 and p. 34.

12 Mark Timmons, 2002, "Motive and Rightness in Kant's Ethical System", in: Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretive Essays*, Oxford: Oxford University Press, ch. 11.

requires (the content of duty), but is a precondition of duty, according to Timmons.

I think Timmons's reading of Kant is right (though I reject ethical internalism). Duty is always available as an incentive to adopt the ends prescribed by the duties of virtue, but one complies with these duties even when one adopts these ends from some other motive. But even if one accepts Timmons's view that for Kant duties of virtue require us only to adopt certain ends, not to adopt them from duty, one might interpret Kant as claiming that to have the relevant virtue one must not only comply with the duty of virtue, but also comply from duty. On this view one's internal act of adopting the required end must be both in accordance with duty and from duty, both right and good. If one complies with this duty from some other motive, one will have done what is required of one, but will not possess the relevant virtue.

Support for this reading is provided by Kant's argument for his account of moral worth, at least on one interpretation of this argument.[13] Kant's view seems to be that a morally good motive must be one that is non-accidentally related to the rightness of the act done from it (when that act is right). Although motives of inclination may lead one to do the right act, when they do, it is just through an accidental alignment of one's interests and circumstances. It is just a matter of luck that the self-interested shop-keeper is honest with all of his customers. He is honest because circumstances are such that he cannot get away with cheating them. If he could, self-interest would lead him to act wrongly. This is not the case with the motive of duty. When I do some right act just because it is right, the rightness of my act is non-accidentally related to my motive for doing it. If I do the right act from a sense of duty, it will be no accident that my act is right.

This argument for Kant's account of morally good motives seems to apply equally to his account of the virtues. It should be no accident that someone who is genuinely beneficent should comply with the duty to make others' ends her own. Indirect egoists would comply with this duty only because they think that doing so will make their own lives go better. But to comply with this duty, they must *really* make others' ends their own – that is, they must somehow get themselves into a state where they have a non-derivative concern for the needs and interests of others. They

13 See Herman, 1993; Marcia Baron, 1995, *Kantian Ethics Almost without Apology*, Ithaca: Cornell University Press; and Philip Stratton-Lake, 2000, *Kant, Duty and Moral Worth*, Routledge: London.

will not have complied with this duty if they care about others only instrumentally. To do this would not be to make others' ends their own, but would be to merely act *as if* they had made others' ends their own. Given that successful ethical egoists would comply fully with the duty of beneficence, and so would come to have a non-derivative concern for the well-being of others, why shouldn't we say that they have the virtue of beneficence?

I think Kant's answer would be: because it is only a matter of luck that indirect egoists have come to have this concern. It is just a lucky coincidence that making others' ends their own makes their own life go better. If circumstances were different, egoists would see no reason to make others' ends their own. This contingency would cast doubt on whether their altruism counts as a genuine virtue in just the way that the accidental relation to rightness casts doubt on the moral worth of motives of inclination.

I would suggest, then, that for Kant, to have the virtue of beneficence is not merely to comply with the imperfect duty of beneficence, but involves compliance from duty. This is not because this duty requires us to make others' ends our own *from* duty. It does not. The motive of duty is not part of the content of this obligation, but is a necessary condition for compliance with it to count as having the virtue of beneficence.

IV. Relating Virtue and the Virtues

So far we have seen that, for Kant, to be virtuous is to be disposed to act from duty, and that to have any of the specific virtues is to act in accordance with one of the various duties of virtue (from duty). Virtuous acts are those that manifest the agent's virtue, i.e., her general disposition to act from duty, but actions done from duty can have moral worth even when they do not realize this disposition (because the agent does not have it).

I said at the beginning that a theory of virtue should show how the various aspects of virtue relate to each other. Kant's theory links virtuous acts with virtue as such. Virtuous acts are related to virtue as the realization of the disposition that constitutes virtue. Furthermore, compliance with a duty of virtue only constitutes having some specific virtue if the agent complies from duty. Such compliance need not be the realization of virtue as such. One might comply with the duty of beneficence from duty, but lack a general disposition to do dutiful acts

from duty. But the more duties of virtue one complies with from duty the more likely it is that one has a virtuous disposition.

There is, however, a striking way in which the various aspects of Kant's account of virtue do not tie up. On his account, an agent's virtue cannot be realized in the exercise of one of her specific virtues. Take beneficence. The virtue of beneficence is exercised when the agent helps someone in need. Although someone who exercises this virtue will have acted in accordance with the imperfect duty of beneficence, there is no duty to *act* beneficently in Kant. The duty of beneficence is not a duty to do a certain type of action (to help others), but is a duty to adopt a maxim of ends. If one adopts this maxim from duty then one has the virtue of beneficence, and specific beneficent acts will realize this specific virtue. But beneficent acts are not morally required, according to Kant.

Since there is no duty to do beneficent acts in Kant's doctrine of virtue, one cannot (knowingly) perform such acts from duty. Since being virtuous is being disposed to act from duty, one cannot realize one's virtue in the performance of beneficent acts (unless one falsely believes that such acts are morally required). If one's moral beliefs are correct, then one cannot do beneficent acts from duty.

This brings us to a problem with the two aspects of Kant's account of virtue. Acts that realize the specific virtue of beneficence cannot realize the agent's virtue and so cannot be virtuous acts, or even morally worthy acts. For virtuous acts are those that realize the agent's disposition to act from duty, and morally worthy acts are those that are done from duty. This is a very odd consequence for any theory of virtue.

It may be thought that Kant could easily get round this problem by maintaining that certain beneficent *acts* as well as the adoption of the maxim of beneficence are morally required. If such acts are required then virtuous agents could do them from duty and so realize their virtue as well as their beneficence in doing them, and everything would tie up nicely.

But this solution creates problems elsewhere. To begin with, the duty to do specific beneficent acts does not fit neatly with Kant's division of duties between perfect and imperfect. In terms of their deontic valence such duties would be imperfect because they prescribe rather than proscribe. But in terms of their content they would be perfect, since they prescribe the performance of certain *actions* rather than the adoption of certain maxims. So it is unclear under what classification such duties would fall in Kant's system.

A further classificatory problem arises in relation to Kant's distinction between duties of right and duties of virtue. The key feature here is that duties of right are coercible, whereas duties of virtue are not. We can compel agents to do certain (external) actions, but we cannot compel them to adopt certain ends. If, therefore, there were a duty to do beneficent acts (in addition to a duty to adopt beneficent ends) this duty would be coercible, and so would be a duty of right rather than of virtue. But the suggestion that we have a duty to do beneficent acts as well as to adopt beneficent ends was meant to repair Kant's doctrine of virtue. This suggestion would hardly help if the relevant duty ended up not being a part of his doctrine of virtue at all.

A more serious problem relates to latitude. Kant's doctrine of virtue aims to capture the idea of merit by introducing latitude into the duty of beneficence. One complies with duty by adopting the maxim of beneficence, but this allows for latitude in how one lives up to one's maxim. This latitude allows us to distinguish differing degrees of virtue. Those who are more strongly committed to this principle have more merit and those who do less in the way of realizing their beneficent ends have less. Latitude is, therefore, important for Kant as a way of accommodating differing degrees of merit in relation to beneficence.

If, however, specific beneficent acts were also required, this latitude would be eroded. For it would then be the case either that every beneficent act would be morally required, or that some would. If all beneficent acts were morally required then clearly latitude would be lost. We would lose this latitude even if the duty to perform beneficent acts were limited, so that only some beneficent acts were required, say, those that do not require great sacrifice, or those that do not conflict with other duties. On this view some beneficent acts are required, and others are not. Of those that are required, there would be no latitude. Of those that were not, the original problem would remain – namely that the manifestation of the virtue of beneficence could not also be the manifestation of virtue as such.

V. Disjunctive Duties

What Kant needs, therefore, is a type of duty that (1) pertains to *actions*, (2) allows for latitude and degrees of virtue, (3) makes it possible for agents knowingly to do the relevant actions from duty, and (4) does not unsettle Kant's distinction between perfect and imperfect duties, and

duties of right and of virtue. The notion of a disjunctive duty meets all four requirements.

A disjunctive duty is one that has disjunctive content: it is a duty to do either A, or B, or C, or… The more disjuncts that figure in its content, the more complex the disjunctive duty will be, though a disjunctive duty may have only two disjuncts, and so may be very simple. The disjuncts of some disjunctive duties will be specific actions. These duties will, therefore, pertain to actions rather than to maxims of actions. So such disjunctive duties meet the first requirement listed above.

But a disjunction of actions is not itself an action. So although this type of duty pertains to actions, such duties do not require that we do some action. If I have a duty to do either A or B, then neither A nor B is obligatory. Each disjunct of the disjunctive obligation is permissible. It is the disjunction itself that is obligatory, not any of its disjuncts.[14]

It is this feature of disjunctive obligations that allows for latitude in acting in accordance with duty. Take the following example of a simple disjunctive obligation. Suppose someone has been injured in a road accident and needs to go to hospital. Suppose I have a disjunctive obligation either to phone for an ambulance, or to take the injured person to hospital myself. The act of phoning for an ambulance and the act of taking the person to hospital myself are each morally permitted, but are not morally required. So I am not obligated to do either act. I am, however, obligated to do either one or the other. Which act I do is up to me. So here I have latitude, but the options from which I am allowed to choose are determined by the content of the disjunctive duty.

But the problem was to get latitude *and* to make it possible to perform some specific beneficent act *from* duty. How could I do one of these acts from duty if neither act is obligatory? Although neither act is obligatory, if I do either act I have acted in accordance with duty. I will have acted in accordance with the disjunctive duty to either call for an ambulance or to take the victim to hospital myself. Since in doing one of these acts I have acted in accordance with duty, we get the surprising result that I can do one of these acts *from* duty even though there is no duty to do that act. If I believe that I ought either to phone for an ambulance or to take the victim to hospital, and decide to phone for an ambulance, I could do this act from this belief. I would then be acting from duty even though the act I chose to do was not obligatory. The notion of a disjunctive duty can, therefore, meet the third requirement

14 W. D. Ross pointed this out in *The Right and the Good* (Ross, 2002, p. 3).

listed above – namely, that it makes it possible for agents knowingly to do the relevant actions from duty.

Do disjunctive duties blur the distinction between perfect and imperfect duties, or between duties of right and duties of virtue? We distinguished imperfect from perfect duties with respect to their deontic valence, their content, coercibility, and merit. Imperfect duties tell us what we must do whereas perfect duties tell us what we must not do. Imperfect duties prescribe the adoption of certain maxims, whereas perfect duties proscribe certain actions. Imperfect duties cannot be coerced, though most perfect duties can. Finally, compliance with imperfect duties is meritorious, whereas compliance with (coercible) perfect duties is not. How do disjunctive duties fit into this schema? Are they perfect, imperfect, or neither perfect nor imperfect duties?

Disjunctive duties, as I have described them, are prescriptive. They tell us that we must do this or that. In this respect then they fall on the imperfect side of the distinction. But in terms of their content they do not fall into either category. Disjunctive duties prescribe disjunctions of actions, and a disjunction of action is neither an action nor a maxim. So their distinctive content means that they cannot be subsumed under Kant's distinction between perfect and imperfect duties. Rather, they constitute a distinct, third type of duty. This does not upset Kant's original distinction (between imperfect and perfect duties). It merely adds to Kant's distinction without problematising it.

Can disjunctive duties be coerced? It seems to me that there is not a simple answer to this question. In respect to beneficence, each disjunctive duty may be summed up as saying that you should do something to help, and this could in principle be coerced. One could, in principle, be punished for failing to do something to help those in need when one is able to help. But two factors conspire to make it practically impossible to enforce compliance with such duties. First, since the disjuncts will be different for each situation, what counts as complying with the duty will be different in different situations. This variability will make it very difficult to enforce by threat of punishment. Second, it will be very difficult to determine what counts as an opportunity to help, or who it is one is able to help. Today there are many people around the world who are in desperate need and for whom help is provided by various charities. Since it is relatively easy to give to such charities, we will always have an opportunity to help very many different people. But it is hardly practical to punish everyone who misses such an opportunity, since just about everyone often misses opportunities to help. So although disjunctive

duties could in principle be backed by coercive laws, such laws would be both too complex and too disruptive to be practical.

Kant held that compliance with imperfect duties is meritorious. We have seen that Kant offers two grounds of merit. The first is a normative ground. Such compliance is meritorious because we do more than what we owe and thus more than what others are entitled to coerce us to do. The second is a non-normative ground. Compliance with imperfect duties warrants merit because in complying we are doing more than we can be coerced to do.

I have said that compliance with disjunctive duties can in principle be coerced, so if such compliance is meritorious, its merit cannot be grounded in its non-coercibility. But this does not mean that compliance with disjunctive duties lacks merit, for there is a further issue of whether we are entitled to coerce compliance. If we are not, then the normative ground of merit would remain in the absence of the non-normative ground, and the merit of such actions would be preserved. I think Kant, following Grotius, would maintain that those in need have no right to force those who could help to aid them. If this is right then Kant would deny that in helping we are doing what is owed and thus would deny that it is legitimate to coerce such actions. If coercion is not legitimate, then compliance is meritorious.

Furthermore, it will be meritorious to the extent that one does more than one may be legitimately forced to do, so degrees of merit can be determined in just the same way as they are for imperfect duties – that is, by the sacrifices involved. If, to go back to my motoring accident example, I chose to phone for an ambulance, this act will warrant some praise. But since the cost to me is pretty minimal, the degree of merit will be rather low. Driving the accident victim to hospital involves a greater cost to me, and so a greater sacrifice. So if I complied with the disjunctive duty by driving the victim to hospital myself, then this act would be more meritorious.

VI. Advantages of Disjunctive Duties

If Kant were to allow disjunctive duties into his account of virtue he could tie up the various aspects of his account. Actions that realized the specific virtues could be done from duty and thus could realize the agent's virtue as such. Disjunctive duties have the further advantage that they introduce strict limits on the latitude we have. Both Kant's earlier

(*Groundwork*) view and his later (*Metaphysics of Moral*) view have implausible implications. On the earlier view the duty of beneficence allows latitude in the interest of inclination. This means that we may not help a stranger in need if we are not inclined to help, so long as we have adopted the maxim of beneficence. This is surely mistaken. I have not done all that morality requires of me if I ignore someone in need, especially if I could easily help.

On Kant's later view latitude is not allowed in the interest of inclination, but in the service of some other maxim of duty. So I might, for instance, benefit a loved one rather than a stranger. But although this is more plausible than Kant's earlier view, it still has implausible consequences. Consider, for example, a situation where I could easily save a stranger from terrible suffering or could confer some minor benefit on a loved one, but could not do both. It seems that Kant's later view allows us latitude to choose to confer the lesser benefit on our loved one over the significantly greater benefit on the stranger. But this seems wrong. It would be wrong to benefit a loved one in such circumstances when one could instead alleviate the terrible suffering of the stranger.[15]

Supplementing his view with disjunctive duties gets around this difficulty, for we may assume that the implausible options in these two cases are not disjuncts of the disjunctive obligation. In both cases the disjuncts would include various ways and degrees of helping the stranger. Some of these ways and degrees would be more demanding, and so more meritorious. Others less. But it would be implausible to think that this disjunctive duty would allow us to ignore the stranger's plight altogether, either because we were disinclined to help, or in order to confer some relatively minor benefit on a loved one. So disjunctive duties impose constraints on legitimate ways of living up to the maxims that duties of virtue require.

Kant's doctrine of virtue needs disjunctive duties. It needs them both to enable agents to manifest their virtue in the realization of some specific virtue, and to place plausible limits on the sort of latitude that imperfect duties introduce. The details of this account would need to be developed, but I think that once disjunctive duties are integrated into Kant's account

15 Baron argues that the duty to perfect oneself means that one can never neglect those in need when one is in a position to help (Baron, 1995, pp. 42 ff). But this duty is also imperfect, and so allows of latitude in the interest of some other maxim of duty. I do not, therefore, see how this duty helps.

of virtue, his account would offer a plausible alternative to Aristotelian-inspired accounts.

Emotions in Kant's Later Moral Philosophy: Honour and the Phenomenology of Moral Value

Elizabeth Anderson

"Love of honour is the constant
companion of virtue." – *Anthropology
from a Practical Point of View*

Those inclined to believe that emotions and feelings lie at the heart of morality are unlikely to turn to Kant for illumination. Kant denied that emotions could offer any foundation for morality.[1] He denied that actions motivated by feeling, even such lovable feelings as compassion, have moral merit [*G* 4: 398]. Rather, whenever we are tempted away from our duty, sensible inclination is the culprit. Hence morality demands that we cultivate "apathy" (the disposition to curb our enthusiasm for any end) and constantly discipline our feelings by reason[2]. Emotions are "pathological"; passionate action is always morally objectionable ([*MM* 6: 408]; *APPV* 155–7).

Although many find these Kantian claims misguided and repugnant, Kant's views about emotions and feelings deserve a closer look. For Kant had a deep insight about the structure of values, which he saw that we grasp at first through our feelings. This is the fact that values appear to us in two dramatically different forms, as appeal and as command [*MM* 6: 379–80]. Appealing values constitute the domain of the good, commanding values the domain of the right. We adopt most ends because they appeal to us. But we adopt other ends because we feel we

1 Immanuel Kant, 1981, *Grounding for the Metaphysics of Morals*, trans. by James Ellington, Indianapolis: Hackett (hereafter abbreviated as *G*), pp. 442–4; Immanuel Kant, 1991, *The Metaphysics of Morals*, trans. by Mary Gregor, Cambridge: Cambridge University Press (hereafter abbreviated as *MM*), p. 376. Note: Pagination for all texts by Kant except for the *Anthropology, Conjectural Beginnings*, and *Lectures on Ethics* refer to the Prussian Academy edition of Kant's complete works.

2 [*MM* 6: 408–9]; Immanuel Kant, 1978, *Anthropology from a Pragmatic Point of View*, trans. by Victor Lyell Dowdell, Carbondale & Edwardsville: Southern Illinois University Press (hereafter abbreviated as *APPV*), p. 158.

must. These ends command; they do not tempt. "It is a lawgiving which constrains, not an allurement, which invites" [*MM* 6: 219]. Kant rightly associated this feeling of necessity with the characteristic phenomenology of moral duty. We do our duties because we have to, not because we want to. We feel that even if our duties don't appeal to us – and often they don't – we would still have to fulfill them, however reluctantly. Often we do them dragging our feet. Commanding values carry authority: they constrain us, and we look up to them. We judge ourselves by their standards. By contrast, appealing values lack authority. They may entice and beseech, coax and wheedle, allure and charm us; they may even ensnare us (as in addictions). But they do not command, for we judge them by standards that are up to us, and so look down upon them.

This distinction between appealing and commanding value is something we *feel.* Kant thought that when we respond to these distinctive forms of value in action, we act with feeling as well. For all his suspicion of feelings, Kant did not repudiate heartfelt action. We must not be misled by the *Groundwork*'s depiction of the cold man of moral virtue into thinking that Kant viewed moral action as motivated by principle alone, without any feeling. Rather, when an agent adopts the moral law as her principle of action, she is moved by distinctively *moral feelings:* most importantly, respect for the moral law and for its source, which is practical reason. Consciousness of the authority of the moral law arouses other moral feelings as well: negative feelings of guilt, shame, and contempt (directed toward those who fail to do their duty), and positive feelings of contentment (moral relief at the vindication of one's moral innocence), pride and honour (directed toward those who manifest moral virtue) [*MM* 6: 399–403]. All these feelings are responses to commanding value.

In this essay I shall stress two themes. First, Kant was not fundamentally hostile to feelings, not even to feelings oriented toward appealing values. His later moral philosophy, especially *The Metaphysics of Morals*, called for the cultivation of many feelings, both moral and nonmoral, and credited the nonmoral feelings with moving humanity toward recognition of the moral law and helping us to follow it. Second, Kant was right to associate the feeling of commanding value with moral rightness, but wrong to ground it *exclusively* in the categorical imperative. The genealogy of commanding value originates in the ethic of honour. Kant was aware of the deep connection between the ethics of honour and his own moral system. Both stress the primacy of pride over self-love, of honour over mere advantage or avoidance of harm. Both stress the

priority of duties to the self over duties to others. Both locate a motive to ethical conduct in a dignified agent's proper regard for himself. Both are sensitive to the dangers of (what passes itself off as) unalloyed benevolence.

Once we understand the honorific origins of Kantian moral responses, we will better appreciate the sources of Kantian resistance to Christian and contractualist ethics. We will also better understand a central difficulty with Kant's applied ethics: the deplorable harshness of its judgments regarding the victims of coercion. Something has gone seriously wrong with an ethic that condemns rape victims as dishonourable for submitting to violence rather than fighting to their deaths. The solution to this problem requires that we reconfigure the principles that inspire our feelings of pride and contempt, honour and dishonour. This is possible, because it has already been achieved to the extent that we have moved from a hierarchical ethic of honour to a more egalitarian ethic of universal respect.

Kantian Moral Psychology: Varieties of Appeal and their Connections to Morality

Kant held that capacities for feeling are capacities for taking pleasure or displeasure in representations of objects [*MM* 6: 211]. Officially, he endorsed the orthodox hedonistic doctrine that pleasures and displeasures vary only in quantity, not quality.[3]

But this doctrine ignores the vast difference in quality of feeling between appealing and commanding value. Kant's later ethical writings reflect a richer, more pluralistic phenomenology of value than even this dichotomy suggests. For example, Kant held that some vicious acts, such as suicide, are horrifying, while others, such as bestiality, are disgusting and nauseating, and still others, such as avarice, are detestable.[4] Each of these feelings responds to commanding value.

We can explain this pluralistic phenomenology of value by considering three factors: (1) how feelings affect and reflect awareness of the orientation and disposition of the body; (2) how they engage the will; (3)

3 Immanuel Kant, 1956, *Critique of Practical Reason*, trans. by Lewis White Beck, Indianapolis: Bobbs-Merrill (hereafter abbreviated as (*CPrR*), pp. 22–3.

4 Immanuel Kant, 1979, *Lectures on Ethics*, trans. by Louis Infield, Indianapolis: Hackett, pp. 150, 159.

how they inspire certain conceptions of the self in relation to the object of feeling. Kant used all three factors to distinguish the varieties of subjective orientation to different types of value. With respect to the body, for instance, he observed that passive entertainment, when experienced for a long time to the exclusion of purposeful occupation, makes one bored and enervated. By contrast, occupation – an active mode of engagement with a purpose – invigorates the body, making one feel alive and aware of one's powers (Kant 1979: 160–1). However, Kant was more interested in the last two factors than in visceral experiences.

Kant thought that there were five basic practical forms of appeal (and aversion): affects, wishes, desires, passions, and aesthetic taste. These can be distinguished by the ways their objects relate to a person's will. Affects (*Affekte*) are momentary unreflective emotional impulses. Anger might make someone lash out at another. In approaching or recoiling from the appealing or repelling object, a person moved by affect does not view his conduct as a means toward an end. He just reacts. Since affective behavior involves neither a representation of an end nor of a means; it is not based on a maxim and is unintentional. Indeed, it blocks the reflection needed to frame maxims ([*MM* 6: 407–8]; *APPV* 155).

One step up from affects are wishes. In wishing, one represents an object – some state of affairs felt as appealing, but not realized – as an imaginary end, without connecting it to the means needed to achieve it (*APPV* 155). One deplores some sad state of affairs, but does nothing to improve it. The end is merely notional and engages only feeling, not will.

Desires occur when our feelings fully engage our will and understanding. When a person conceives of an appealing state as her end, which she shall bring about by undertaking the necessary means, she desires the end. Desires are embodied in maxims. They involve a representation of the agent bringing about her end through that representation [*MM* 6: 211].

Passions (*Leidenschaften*) are desires raised to the level of obsession [*MM* 6: 407–8]. In passion, feelings are incorporated into lasting maxims that govern conduct monomaniacally. Greed and miserliness entrench the desire for money, making it difficult to steer another course (*APPV* 156). Passions are calculating and willful, but lack the free play characteristic of ordinary desire, whereby the will adopts and casts off ends as other objects of appeal are imagined.

Aesthetic taste is the pleasure we take in the representation of beautiful objects. In contrast with wish, desire, and passion, which takes pleasure in the thought of some state being realized, aesthetic

appreciation takes pleasure in the mere thought of an object apart from its existence. Thus, this feeling is purely contemplative, detached from desire and will. Unlike affects, aesthetic feeling does not move the body and requires a tranquil state of mind. It is steady, survives reflection, and is universal rather than idiosyncratic [*MM* 6: 212–3].

These different forms of appeal alter the awareness of the subject in relation to the object of feeling. We lose ourselves and fall apart when we are overcome by affects. To recover from their influence, we must get hold of ourselves again. This is why Kant considered affects to be weak and childish [*MM* 6: 208]. By contrast, desire gives people an inchoate sense of their freedom in relation to a world of choices. This arises from the fact that the spontaneous exercise of imagination and understanding gives rise to new, nonnatural objects of appeal, including desires for luxury, refinement, and culture. Imagination makes desire run riot, expanding the field of perceived choices and highlighting the contrast between humans and animals, whose desires are tied by natural instinct to fixed ends.[5] Desire also gives people their first sense of themselves as ends in relation to a world of mere things, which are seen to exist merely to serve them:

> The first time [man] said to the sheep, "Nature did not give thee the pelt thou wearest for thyself, but for me," stripped him of it and put it on himself, he perceived a prerogative, that he, by virtue of his nature, was above all animals, which he now considered…as the means and instruments left to his will for the accomplishment of his purposes at pleasure (Kant 1798: 324).

In pursuing consciously chosen desires, people see themselves as elevated above mere things.

This is why succumbing to passion feels degrading. The desperation of a drug addict for another high turns the person into the servant of her passion, when she should be master of it. No wonder Kant condemned passions as pathological. "Who wants to put himself in chains, when he can be free?" (*APPV* 157). Aesthetic feeling, by contrast, gives people a sense that some things are *not* just for their use; they see that they can love and find value in things outside themselves and the satisfaction of personal desires. Aesthetic feeling, however intensely felt, cannot enslave, because it does not move us to action.

---

5 Immanuel Kant, 1798, "Conjectural Beginning of the History of Mankind", in: *Essays and Treatises on Moral, Political, and Various Philosophical Subjects*, vol. 1, London: William Richardson, pp. 321–3.

Kant argued that no form of appealing value can ground morality, because it cannot give rise to obligation, either subjectively, as the feeling of command, or objectively, in an authoritative principle. One does not have to be a Kantian to appreciate this point. John Dewey agreed with Kant that it undermined consequentialist definitions of the right as what maximizes the good. The good, Dewey argued, is what appeals to desire, but the right cannot be simply what maximizes desire fulfillment:

> [A]s an idea, "right" introduces an element which is quite outside that of the good. This element is that of exaction, demand… The Good is that which attracts; the Right is that which asserts that we ought to be drawn by some object whether we are naturally attracted to it or not.[6]
>
> We find ourselves subject to authority, under law, bearers of responsibilities which we did not choose and which we must meet. There is a strain of authority and obligation in morality which is not … reducible to the conception of the good as satisfaction, even reasonable satisfaction, of desire (Dewey and Tufts 1932: 214).

Besides attacking theories of the right grounded in appealing value as a general matter, Kant also criticized particular theories of morality for their dependence on particular types of appealing value. He criticized the "monkish" ethic of self-denying asceticism as fanatical – that is, based on a passion (Kant 1979: 158). He criticized the sentimental ethic of compassion or impersonal benevolence for smugly wallowing in idle wishing for others. If I feel toward this person who needs my help only what I feel impersonally to humans generally, my attraction to their well-being would be so weak as to fall short of engaging the will [*MM* 6: 451–2]. A person moved by such merely sentimental love for humanity thinks: "wouldn't it be nice if all the hungry were fed," and wrings her hands over reports of starving children in Niger. She doesn't reflect on what it would take to eliminate starvation in Niger, much less take any steps to achieve this end. Yet she pats herself on the back for merely wishing them well (Kant 1979: 200).

Although Kant insisted that appealing values cannot function either as the ground of morality (principles of right) or as the fundamental motivation to be moral, his position should not be construed as a repudiation of feelings of appeal, even from a moral point of view. He thought they had two critical instrumental roles to play on behalf of morality. First, they are stepping stones to morality in the historical

6 John Dewey/James Tufts, 1932, *Ethics*, New York: H. Holt and Company, p. 216.

development of human civilization. Second, they assist moral feelings in helping people do their duty.

Kant believed that although the moral law can be known *a priori*, humans have not been conscious of it from the start, much less known how to arrange their institutions so as to realize its demands. Coming to universal awareness of the moral law and its demands is an historical achievement of an enlightened age, which he saw as yet to come. The process of arriving at an awareness that we *need* the moral law and that it demands universal respect for every rational being involves reflection on the predicament that appealing values place us in. As humans expand their imaginations and understanding, they frame novel desires that place increasing demands not only on their production technologies but on their social arrangements. Desires run ahead of the economic and social means available to satisfy them. They entrench themselves in passions for luxury, esteem, and power. This puts people into conflict, competing over the means to satisfy their desires and for social recognition. Our "unsocial sociability" – our need to engage in social cooperation to satisfy antagonistic passions – points us beyond our desires to the need for social arrangements that can manage conflict in a way that all can accept – that is, in a way consistent with everyone's equal freedom.[7]

Why *equal* freedom? As noted, awareness of their proliferating desires makes people think of the world of things as mere means, things that exist for their sakes. In this way, all come to think of themselves as ends, not means. So they resist being treated as mere means by others, and demand recognition from others that they are ends. People feel free to take the sheep's pelt, but they have the inchoate thought that treating humans like sheep would transgress against their claim, which, springing from the same sources as their own, has equal validity (Kant 1798: 325). People do not fully develop and credit this thought until, exhausted by warfare, they are driven to seek peaceful ways to resolve their conflicts. Meanwhile, their passions for luxury, power, and esteem lead them to develop the arts and sciences, talents and social graces (by which we cultivate the appearance of respect for others, whatever our inner feelings may be) (Kant 1979: 175–6). These, in turn, equip them to design laws and social institutions that follow the moral law.

---

7 Immanuel Kant, 1963, "Idea for a Universal History from a Cosmopolitan Point of View", in: *On History*, trans. by Lewis White Beck, Bobbs-Merrill, http://www.marxists.org/reference/subject/ethics/kant/universal-history.htm (10. October 2005).

Thus, the pursuit of appealing values prods us to consciousness of the moral law and its demands. Some of them also assist us in following the moral law, and so should be cultivated on that account. Morality requires that we love and respect everyone in a practical sense: that we make it our policy to promote their happiness and give them their due [*MM* 6: 448]. Moral feelings induce us to adopt this policy. But they are not always strong enough to overcome temptations to falter in our execution of it, and do not automatically enable us to see what it demands, especially when this requires collective action in constructing just institutions. Feelings of sympathy for others, which spread by association or "infection" rather than principle, often step in to fill the gap between what we owe people and what our social arrangements give them. Sympathy is "one of the impulses that nature has implanted in us to do what the representation of duty alone would not accomplish" [*MM* 6: 457]. Although charity is no substitute for justice, it helps us render what we owe people. Moreover, once we make beneficence our steady practice, compassion follows. So, although feeling compassion cannot be a direct duty, since it cannot be willed, we do have a duty to *cultivate* compassion as an aid to doing our duty.

For the same reason, we have duties to cultivate our taste for other appealing values. The social graces – affability, etiquette, tolerance – "promote the feeling for virtue itself by a striving to bring this illusion as near as possible to the truth. They make virtue fashionable" – that is, appealing ([*MM* 6: 473–4]; see also *APPV* 37). Aesthetic taste awakens a recognition that we can love things without using them. This nonmoral feeling prepares us for morality, which requires that we value human beings in ways apart from using them. Wanton destruction of beautiful objects is wrong because it degrades our capacity for aesthetic feelings, which aid morality [*MM* 6: 443]. We even have a duty to cultivate cheerfulness in doing our duties and bearing our burdens ([*MM* 6: 485]; Kant 1979: 146). These are ways to adorn moral demands with forms of appeal that are not inherent to them. We thereby enlist the nonmoral motivations in the service of morality.

Many people are likely to find even this more generous assessment of feelings of appeal in Kant's later moral philosophy to be too stinting. I do not find it too stinting within its very narrow focus. Kant was almost exclusively preoccupied with the right rather than the good. He correctly held that feelings oriented toward appealing values can provide no basis for the right: they cannot explain why the right has commanding authority. He also correctly viewed these feelings to be only contingently

connected to the right, sometimes helping us do our duty, often not. Kant's focus on considerations of right entailed that his assessments of feelings of appeal were limited to considering how they contribute to or impede our fulfillment of our duties. Nothing in Kant's philosophy precludes other sorts of appraisals of feeling – for instance, considering whether particular feelings are lovable, beautiful, or admirable. We love people who are sweethearted, find graciousness beautiful, and admire people with fine aesthetic taste. These favourable responses to feeling are merited. But they have nothing to do with moral duty.

In Search of Moral Feeling: The Genealogy of Commanding Value in the Ethics of Honour

Consider now the phenomenology of commanding value. From a naturalistic point of view, we can explain the feeling that we *must* do something as arising from our experience of being ordered to do so by an acknowledged authority. Commanding value, then, is essentially possessed by persons, not things or states of affairs. "The constraining (binding) subject must be a person" [*MM* 6: 442]. Confronted with an authoritative command, we feel respect for it and for the person who issues it. The reluctance with which we obey commands follows from the fact that they do not spring from our own desires, but from another's will. While under orders, we feel that our performance will be judged. Hence, we feel humbled before authoritative commands: we measure ourselves by their standards. This experience poses both a risk and an opportunity. The risk is that we won't measure up to the authoritative demands. If we don't, we feel guilt – the feeling that we are accountable to our commander, stand accused of failure, and cannot acquit ourselves of the charge. We also feel shame – the feeling of personal inadequacy that makes us want to hide from judgment. The opportunity is that we will measure up to the demands made upon us. If we do, we feel relief and vindication – the contentment that arises from being acquitted of charges [*MM* 6: 391]. We also feel self-esteem or pride. We hold our heads high, our backs erect, full of a sense of our own honour and rectitude.

Kant interpreted all of these feelings – of respect, conscience (the internalized judge), humility, guilt, shame, vindication, and personal honour – as our modes of response to the authority of the moral law and recognition of our varied relations to it [*MM* 6: 399–403]. They are the

subjective conditions of our ability to be receptive to the commands of duty and motivated to follow them. Without them, we wouldn't feel moral demands to be compelling. The feelings themselves are natural endowments. "[T]he pain a man feels from the pangs of conscience has a moral source, [but] it is still a natural effect, like grief, fear, or any other state of suffering" [*MM* 6: 197]. Yet Kant thought they responded only to the demands of reason – that is, the moral law. "Consciousness of [these feelings]…can…only follow from consciousness of a moral law, as the effect this has on the mind" [*MM* 6: 399].

The phenomenology of commanding value that Kant identified is not peculiar to Kantian morality. Several distinct systems of morality, each tying rightness to commanding authority, draw from the suite of commanding feelings just described. But they harness somewhat different feelings, depending on the content of moral duties they prescribe and the ways they represent moral agents as bound by them. Recall that feelings can be qualitatively distinguished by three factors: (1) our consciousness of their effects on the body, (2) their felt relations to the will, and (3) the conceptions of the self in relation to the object of feeling that they inspire. The most distinctive feature of Kant's phenomenology of right concerns the third element. Specifically, it concerns the *pride* or *self-esteem* people feel when they understand themselves to be subject to and executors of the moral law. For Kant, to feel oneself as obligated by the moral law essentially involves a sense of oneself as a dignified or honourable agent. Consciousness that one is worthy of being entrusted with the execution of the awesome responsibilities of upholding the moral law fills one with a sense of self-worth and thereby produces "an exaltation in [one's] soul" [*MM* 6: 483]. Hence, the moral agent meets others eye-to-eye, cowering before no one, confident in his claim to a form of personal worth that commands the esteem of all. "The human being believes he feels his personal worth through [obeying the moral law] alone" [*G* 4: 450].

The feeling in question is not respect, which Kant thought was owed to all rational beings, regardless of their conduct or state of character [*MM* 6: 463]. Respect for a person is what moves people to constrain their actions in accord with that person's rights. The feeling is rather esteem or honour, which, when felt by a person toward herself, is pride. What makes this feeling moral (a response to commanding value) is its link to the person's fulfillment of morally binding commands. Esteem and pride have to be earned through virtuous conduct and can be lost through vicious conduct. Thus, Kant argued that the love of honour

makes one resist lying and other vices that reflect a lack of moral character, for to indulge these vices would make one an object of contempt [*MM* 6: 420].

Not every moral system treats moral self-esteem as a matter of central importance. Those that do have distinctive features. The supreme importance of the sense of personal worth in binding agents to the moral law feeds back into the content of the moral law itself. Not every command can be followed consistent with preserving the dignity of the agent who carries it out.

To see the importance of this point, compare Kant's conception of commanding value with that in three other systems of right: divine command morality, contractualism, and the ethic of honour. These four systems are alike in taking the contrast between appealing value and commanding value seriously. They all refuse to reduce the latter to the former, as consequentialist theories tend to do. They all recognize that the source of commanding value, of principles of right, must be a person, not mere things or states of affairs. They differ most fundamentally in their conception of the source of morally binding commands: in divine command morality, it is God; in contractualism, it is everyone affected by one's actions; in the ethic of honour, it is those who occupy (or deserve to occupy) leadership roles; in Kantian ethics, it is the moral agent herself. Kant's objections to divine command morality and contractualism are based on their failure to protect the dignity or proper pride of the moral agent. In focusing solely on duties to others, they neglect duties to the self and thereby require or permit agents to engage in self-debasing actions. Kant's system shares with the ethic of honour a deep concern with the dignity of the moral agent, both in the content of morality (via duties to the self) and in inspiring agents to follow the moral law. This affinity reflects the fact that Kant's ethic draws from the older ethic of honour. Kant's great innovation was to reconfigure the relationship of respect to esteem. Where the ethic of honour made respect for an agent dependent on his esteemability, Kant made esteem dependent on the respect the agent showed for everyone – *including herself.*

Consider the place of moral pride or self-esteem in each of the four systems of right, starting with Christian divine command morality. It represents the source of moral commands as infinitely superior to those to whom the commands are issued. Moral agents are comparatively worthless. Consciousness of one's relative worthlessness and powerlessness is critical to binding the agent to God's commands. God is awesome and terrifying; so people must kneel and tremble before God and dare not

question what he orders, no matter how terrible the deed commanded. Abraham must sacrifice his own son, if that is what God commands. Consciousness of one's own relative worthlessness is also critical to the love some religious subjects have for God. No one deserves salvation, for all are sinners. People therefore beg for God's mercy, his gratuitous love for human beings, to bring salvation. The love of God is thus a form of gratitude, which, as Kant stressed, situates the beneficiary in an inferior relation to the benefactor, as a degraded and humiliated subject of pity ([*MM* 6: 455, 459]; Kant 1979: 118–9).

Kant was appalled by divine command morality. He thought that the submissive posture of blind obedience it requires of moral agents was grossly inconsistent with morality. Abraham *must* challenge the purported authority of the voice that tells him to murder his own son, because he knows by his own reason that this act is wrong.[8] The conscientious moral agent must therefore always reserve judgment for himself when faced with any command.

Divine command morality's stress on the lowly status of human beings leads to other objectionable practices. Without a concept of human dignity to constrain them, some believers in divine command morality are prone to religious fanaticism, "the most ungodly of all passions…because it makes man think that under the cloak of piety he can do all manner of things" (Kant 1979: 146). To be moved by passion is always undignified; the moral transgression is made worse by the conviction that, since the value of human beings is conditional on God's saying so, we can ignore their complaints whenever a voice we think is God's tells us to violate their rights. In addition, without a sense of personal dignity, the fanatically devout engage in shameful forms of public prayer (Kant 1979: 115), disgracefully prostrating themselves before their image of God [*MM* 6: 436]. They adopt degrading, monkish practices of self-mortification and self-denial that violate one's duties to oneself (Kant 1979: 158; *APPV* 191).

Contractualism, which locates the source of authoritative moral claims in other people, bears a closer affinity to Kant's ethics. With Kant, it stresses the equal standing of human beings and limits the possible content of authoritative commands of right to what is consistent with the reasonable reciprocal claims of everyone. However, it limits the objections that can be made against an act to the claims of the *particular persons,*

8 Immanuel Kant, 1960, *Religion Within the Limits of Reason Alone*, trans. and ed. by Theodore Greene and Hoyt Hudson, New York: Harper & Row, p. 175.

other than the agent, who are affected by them.[9] Kant thought that acts could be wrong on two additional grounds: because they assault humanity in general, and because they violate the dignity of the agent. Spreading scandalous gossip, if it is true and does not violate privacy (say, because the facts reported are a matter of public record), does not violate the rights of the scandalized person. Yet Kant still thought it wrong because it expresses and promotes misanthropy and "diminishes respect for humanity as such" [*MM* 6: 466]. More controversially, Kant argued that a cheater, in adopting a maxim of cheating, has waived his right not to be cheated. If his victims cheat him in return, they do *him* no wrong. Their wrong consists rather in an attack on humanity (Kant 1979: 230).

Kant's more central objection to contractualism was that it neglects duties to oneself. Contractualism permits degrading contracts between people, so long as the parties give their consent. Thus, it allows contracts into prostitution, concubinage, and servitude. Kant objected to all such contracts, as violations of an agent's duties to protect her own dignity [*MM* 6: 278–9, 283]. Similarly, if the persons harmed by an otherwise morally problematic act voluntarily relinquish their claims, contractualism withdraws its objections to the act. Kant insisted that people have a duty to themselves to press their rightful claims ([*MM* 6: 436]; Kant 1979: 214). Servility – allowing others to "tread with impunity" on one's rights, or selling one's dignity for mere advantage – violates duties to the self [*MM* 6: 435–7]. Kant devoted pages to explaining why suicide, drunkenness, gluttony, and personally defiling acts of lust violate the dignity of the self (Kant 1979: 147–171; [*MM* 6: 421–8]). He even argued that lying is more a violation of duty to oneself than to others, because it dishonours the self [*MM* 6: 429]. Contractualism has no place for duties to the self. Kant claimed that duties to the self take priority over duties to others (Kant 1979: 118). One must not make oneself contemptible.

This contrast in the content of Kantian and contractualist moral systems follows from the different moral feelings they deploy in the service of commanding value. The fundamental contractualist moral attitude is reasonability. It directs us to constrain our pursuit of appealing values out of regard for the claims of others. It is an essentially other-regarding motive. Contractualism relies on prudence, not moral feeling, to ensure that people don't accept terms disadvantageous to themselves.

9 Thomas Scanlon, 1998, *What We Owe to Each Other*, Cambridge, Mass.: Harvard University Press.

The fundamental Kantian moral feelings are respect and moral esteem. The self-regarding feelings of prudence, self-respect, and moral self-esteem orient the self toward the world perceived in very different normative terms. Prudence is concerned with the benefits, advantages, and harms that accrue to the self. Self-respect and moral self-esteem are concerned with personal dignity and worth – the value of the self, not the self's circumstances.

The contrast between the commanding feelings deployed by contractualist and Kantian ethics is particularly stark with respect to moral failings. If a contractualist fails to live up to legitimate moral demands, she may feel guilty. But she will not loathe herself. By contrast, a Kantian agent is moved by a much richer and darker set of feelings toward herself upon contemplating her failures to do her duty to others *or to herself:* shame, loathing, contempt, disgust, feelings of defilement, degradation, nausea, even horror. Kant's discussions of our proper responses to violations of duties to the self are dripping with such powerful moral feelings ([*MM* 6: 423, 425, 429, 435]; Kant 1979: 150–1, 159, 163, 170). Feelings of loathing, contempt, and disgust for persons who have failed to live up to valid commands are alien to the businesslike ethic of contractualism, which is oriented to advantage and benefit rather than inner self-worth.

These feelings *are* at home in the noble and soldierly ethic of honour. This ethic locates commanding value in persons who occupy, or are deemed worthy to occupy, positions of command in a hierarchical society – patriarchal heads of families, tribal and clan leaders, military officers, monarchs, nobles, feudal lords. Honour denotes a superiority of status, recognition, or esteem that those in command enjoy over those who are commanded. Those who are commanded are regarded as lowly – proper objects of contempt. Those who legitimately command are superior – proper objects of honour and respect. To be honourable entitles one to rights. The lowly, by contrast, are not entitled to rights against their superiors, although the honourable are considered virtuous for showing magnanimity, for condescending to treat the interests of their inferiors with consideration.

The ethics of honour is the original home of both the inspirational and the darker moral feelings animating Kant's ethics – feelings of dignity, worth, and honour, on the one side, and contempt, disgust, and degradation, on the other. Kantian ethics inherits several of its features: its stress on considerations of honour over advantage, on duties to the self, and on the sense of personal honour or self-esteem as a motive

binding people to moral demands. Yet in each case, Kantian ethics transforms these features in harnessing them to the service of universal and egalitarian values.

Consider first how the ethics of honour embodies these features. It grants honourable men certain rights and privileges of rank, conditionally on their obeying a code of honour. The code requires them to be willing to sacrifice their interests for the sake of their honour. This is most clear in the case of soldiers, who are obligated to face death in battle rather than run away to save their lives. In the ethic of honour, it is better to die honourably than to live dishonourably. This is an example of what I have elsewhere denoted the *hierarchical incommensurability* of values.[10] Some values have a price; others are beyond all price. The values that have a price are the appealing values, including life, health, security, comfort, and luxury. Commanding values such as personal dignity or honour are beyond all price, in that they must never be deliberately traded for appealing values [*MM* 6: 434–5]. The latter must be sacrificed for the former in conflict cases. This entails that pride must trump self-love as the self-regarding motive binding people to their duties to themselves. People show themselves to be unworthy of superior rank, and hence of esteem and respect (rights) precisely by exhibiting a willingness to trade their claim to esteem for the sake of their self-interest. Thus, servile dependents deserve their contemptible status by willingly trading their personal independence in return for subsistence. A man of honourable rank who does likewise forfeits his claim to respect and superior standing. Call this the *forfeiture principle.*

Because of the supreme importance it places on protecting one's standing before others, the ethics of honour places great stress on duties to the self. The man of honour must not do anything that would compromise his standing. To do so is degrading, in the literal sense of making one deserve and possibly end up with a lower social rank. Thus, the ethic of honour treats as duties to the self what we would naturally classify as duties to others. The man of honour is a man of his word, not out of regard for the interests of those to whom he has made promises, but out of regard for his own worth, which would be defiled by a lie. We can trust the man of honour to do his duty to us because we are confident

10 Elizabeth Anderson, 1995, "Practical Reason and Incommensurable Goods", in: Ruth Chang (ed.), *Incommensurability and Value*, Cambridge, Mass.: Harvard University Press.

that his sense of self-esteem would prevent him from sacrificing his precious honour for mere advantage.

The ethic of honour thereby mobilizes feelings of self-worth in the service of moral duty. The sense of oneself as embodying commanding value makes one recoil in shame and disgust from dishonourable conduct, which is always motivated by some lower, contemptible, merely appealing motives of the self – animal appetites, childish emotions, servile self-interests, wanton passions. To allow such motives to override one's sense of honour degrades one to the status of those creatures to whom these motives are proper – animals, children, servants, wantons – and who for that reason lack the qualities needed to be eligible for command. Even to rub shoulders with these beings of lower rank is to risk defilement. To reveal one's own needs and suffering to others is also to put one's standing at risk, by courting pity (a condescending feeling) toward oneself, and dependence on others' beneficence, which is degrading. Maintenance of personal dignity thus requires that honourable men keep their distance from one another, and hide their woes behind a cheerful demeanour.

That Kant's moral philosophy emerged from the ethics of honour explains some of its otherwise puzzling features. In condemning servility, Kant endorsed the forfeiture principle: "He who makes himself a worm cannot complain afterwards if others step on him" [*MM* 6: 437]. He also held that duties to self should take priority over duties to others (Kant 1979: 118). He agreed with the ethics of honour that a person who lacks a fierce sense of her own self-worth cannot be trusted to do her duties to others (Kant 1979: 118). (The self-effacing dutiful wife was not a salient counterexample to Kant, a lifelong bachelor.) He argued that the moral education of youth should stress the shamefulness of immoral action to the self, rather than its harmfulness to the self or others [*MM* 6: 483]. Thus, teachers and parents should cultivate the sense of honour in their wards as an aide to morality. The love of honour – the desire not to appear contemptible in one's own eyes or others – is the foundation of duties to the self (Kant 1979: 124–5). "The love of honour is the constant companion to virtue" (*APPV* 163). To be guided by motives responsive to commanding value makes one merit esteem and pride, a sense of superiority to those driven by feelings attuned to appealing value. To allow oneself to be moved by the lower motives is to "throw oneself away" (Kant 1979: 118–9) – that is, to forfeit one's claim to the superior rank of a dignified agent. This explains why Kant believed that sexual acts such as masturbation and bestiality are degrading, reducing one to the status of animals, and hence a violation of duties to the self (Kant 1979:

170; [*MM* 6: 425]). It explains why he thought prostitution to be a violation of a duty to the self: in prostituting herself, a woman allows herself to be used as the mere object of another's lust, only to be "cast aside as one casts away a lemon which has been sucked dry" once that lust is satisfied (Kant 1979: 163). She thereby trades her standing as a person commanding dignity for the sake of mere advantage or subsistence. Kant also held that dignified people need to keep their distance from one another, and have a duty to bear their woes with cheer and fortitude [*MM* 6: 449, 484–5].

All of these features of Kant's moral philosophy originate in the ethic of honour. Yet Kant radically transformed this ethic in the service of universal and egalitarian values. The ethic of honour reserves respect, the status of being a bearer of commanding value, of rights that constrain what others may do to them, exclusively to people of superior social rank, to people who merit esteem. Kant's ethic universalizes respectful standing to all rational agents, who are thereby entitled to rights. This requires that one be able to respond to other persons as bearing commanding value – the only kind of value that can place categorical constraints on one's own appealing motives – even though they do not have high social rank and indeed have not even done anything to earn this respect. Yet Kantian ethics cannot do without any notion of superior and inferior beings. Without such status distinctions the counterpart feeling to respect – contempt – has no ground. Kant's solution was to displace rankings of respect from the social hierarchy and project it instead on the order of nature. One human is not superior in his claim to respect to any other, but humans as a species are superior in rank to animals. Animals are therefore not entitled to respect or rights of any sort (although we may have reason to treat them kindly, out of our duty to human beings to avoid cruel dispositions that may tempt us to abuse others) ([*MM* 6: 443]; Kant 1979: 239–41). People get their sense of incommensurably superior worth in part from their felt superiority to animals and things, who are regarded as mere objects of desire, as objects of appealing rather than commanding value (Kant 1798: 324).

In universalizing human claims to respect, Kant had to transform respect from an essentially positional and relative value to an absolute value. In the ethics of honour, respect comes in degrees, depending on one's social rank. One might climb up the ladder of social recognition, but only relative to others. The pyramidal social structure guarantees that few can command the highest respect. In the egalitarian ethics of respect, all command an equal respect from everyone else, which requires that all

are entitled to equal reciprocal rights. This requires that one detach claims to respect from measures of worth that one may have relative to other people, and ground it in something all members of society share – their humanity.

The upshot of this transformation is that Kantian ethics reverses the relations of respect and esteem. In the ethics of honour, one lays claim to respect (rights) in virtue of meriting the *esteem* of others. In Kantian ethics, one lays claim to respect in virtue of the humanity one shares with everyone else, and to moral esteem in virtue of taking respect for one's own and others' humanity as one's end. This is a revolutionary development in human morality.

The Problematic Legacy of the Ethic of Honour in Kant's Moral Philosophy

Kant's moral revolution is based on its reconfiguration of the feelings of respect and moral esteem. By tying claims to esteem to active respect for others, Kant may appear to have superceded the disturbing features of the honour ethic, which views the mass of human beings with contempt. But Kant's applied moral philosophy reflects a transitional stage from the ethics of honour to the ethics of respect, in which some of the hierarchical features of the honour ethic persist. To the extent that people's dignity and standing before others still hangs on the older ethic of honour, Kant argues that they may or must heed the honour ethic's demands. We can see this in three aspects of Kant's applied ethics: his insistence that virtuous agents avoid accepting gifts, his surprising sympathy to honour killing, and his condemnation of rape victims for submitting to rape rather than fighting to their deaths.

Consider first the ethics of gifts. Kant repeatedly stressed the moral perils of accepting gifts. To accept another's help is to degrade oneself in relation to the giver. Receiving help gives rise to an obligation of gratitude that can never be fully discharged. For even if one reciprocates, the giver still enjoys a superior standing for being the *first* to have given. Thus, the morally virtuous agent, who has a duty to himself to protect his own dignity, must avoid accepting gifts from others except when his neediness absolutely requires it (Kant 1979: 118–9, 204–5, 218, 222; [*MM* 6: 455, 458–9, 471]).

This way of thinking about gifts is a legacy of the honour ethic. In societies governed by this ethic, gift-giving is a primary way for superiors to acquire dependents who are obliged to obey and honour them.[11] Even loans are subject to the hierarchical logic of gifts. Kant recognized that such an economic system, in which masses of people must trade their dignity to gain access to the means of subsistence, is unjust. He argued that in reality, the benevolence that the rich show to the poor hardly makes up for an unjust economic system that puts masses of people in a state of permanent neediness (Kant 1979: 211, 236). Hence rich benefactors are not entitled to preen themselves over their charitable works or consider themselves superior to their benefactors [*MM* 6: 454]. Yet Kant recognized that people did not see things that way. They were too wedded to the honour-based logic of gift-giving, and so unfairly looked down upon impoverished recipients of charity. Given this fact, Kant insisted that benefactors must do what they can to protect the dignity of recipients, by representing what they give as what recipients are owed, and by giving anonymously if possible [*MM* 6: 448–9, 453].

This makes sense within the terms of an ethic of universal respect. Such an ethic would aim at detaching the rules determining access to subsistence from gift-giving, and allocate goods on the basis of impersonal and universal principles by which all could retain their dignity.[12] But as long as the distributive rules are still unjustly bound up in the logic of honour, Kant argued that the duty to the self of maintaining one's dignity demanded that one refuse gifts and even unsecured loans. Thus, where the flawed honour ethic still prevails, Kant insisted that people accommodate themselves to some of its demands.

A similar point arises with respect to honour killing. Kant considered two cases of honour killing: the soldier who kills someone who has insulted him in a duel, and the unwed mother who kills her baby to spare herself the disgrace of having an illegitimate child. Within the honour ethic, both killings are required to redeem the agent's honour. Since the honour ethic values honour above life, it sees no moral transgression here. Kant's ethic of universal respect places the right to life and the duty to protect it prior to a person's interest in defending his or her honour. Hence both killings are "crimes deserving of death" [*MM* 6: 335). Yet

11 Marcel Mauss, 1967, *The Gift*, I. Cunnison (trans.), New York: W. W. Norton.

12 Elizabeth Anderson, 2004, "Ethical Assumptions of Economic Theory: Some Lessons from the History of Credit and Bankruptcy", in: *Ethical Theory and Moral Practice* vol. 7, pp. 347–60.

Kant flinched from the conclusion that the state ought to punish these killings by death, although he held in all other cases that murder must be so punished. For in a society in which the honour ethic still held sway, the law cannot wipe out its influence by declaring that killing someone for honour is unjustified. So it must accommodate itself to the demands of the honour ethic, lest the law be discredited in the eyes of the people and forfeit its power to move the people beyond their "barbarous" moral condition, which is the product of their flawed social relations [*MM* 6: 336–7]. Kant thus suggested that the ultimate solution to honour killing must lie in a transformation of social relations rather than by demanding that each individual on his or her own resist the demands of a still-intact honour ethic.

Kant's accommodations to the honour ethic do not merely reflect the circumstances of imperfect compliance theory, in which the demands of morality must be determined in light of the fact that not everyone is living up to them. They reflect the fact that Kant's ethic of personal dignity emerges out of the ethic of honour that it has not entirely left behind. This is revealed most disturbingly in Kant's argument that rape victims behave dishonourably in succumbing to rape rather than resisting to their deaths:

> If a man cannot preserve his life except by dishonouring his humanity, he ought rather to sacrifice it … if he can no longer live honourably, he cannot live at all; his moral life is at an end … [because] it is no longer in keeping with the dignity of humanity… Thus it is far better to die honoured and respected than to prolong one's life for a few years by a disgraceful act and go on living like a rogue. If, for instance, a woman cannot preserve her life any longer except by surrendering her person to the will of another, she is bound to give up her life rather than dishonour humanity in her own person, which is what she would be doing in giving herself up as a thing to the will of another (Kant 1979: 156).[13]

Kant's view of rape victims is an application of his general principle that people must not put their dignity up for sale or otherwise sacrifice it for the sake of mere advantage, even for the advantage of life itself. Degrading contracts are forbidden, because they treat as having a mere market price what is of incomparable worth [*MM* 6: 283, 330]. On the same ground, Kant argued that "if a man offer his body for profit for the

13 See also Kant 1979: 150 ("I should endeavor to preserve my own life only so far as I am worthy to live.") and Kant 1979: 152 ("It is better to sacrifice one's life than one's morality. To live is not a necessity; but to live honourably while one's life lasts is a necessity.").

sport of others – if, for instance, he agrees in return for a few pints of beer to be knocked about – he throws himself away" (Kant 1979: 119). Prostitution is also forbidden, even if this is the only way a person has "to save parents and friends from death" (Kant 1979: 119; see also [*MM* 6: 278]). If a woman may not sacrifice her dignity in prostitution in order to save another's life, she may not sacrifice her dignity in submitting to rape to save her own life.

Kant's treatment of life/honour tradeoffs is inconsistent. A consistent morality of universal respect would view one's claims to honour as dependent on showing respect for humanity and hence on the duty to protect human life. So there can be no honour in killing a person for the sake of avoiding disgrace. Kant applied this principle in morally condemning Lucretia's suicide after she was raped (Kant 1979: 149–50), the soldier's killing of his slanderer in a duel, and the unwed mother's infanticide. So why did he insist that the rape victim should fight to her death rather than submit to rape?

Perhaps Kant thought that once a person is disgraced, killing cannot redeem the dishonour. Disgrace can only be avoided *ex ante*; once one's honour has been stained nothing can remove it. But this does not seem to be true for the soldier who proves his bravery by risking his life in a duel. Perhaps he was leaning on the distinction between killing and letting die. While it is wrong to deliberately kill oneself to protect one's honour, putting oneself in a position where someone else may kill one may be justified.[14] In the same lecture that condemned Lucretia, Kant observed that a soldier who bravely faces death at the hands of the enemy does not commit suicide (Kant 1979: 150). Kant thus appears to be drawing an analogy between the soldier's and the prospective rape victim's situations: both rightly protect their honour in fighting to their deaths; neither can be blamed for the fact that their assailant kills them. Although they "allow" themselves to die at the hands of their enemy, they do not murder themselves.

Closer scrutiny of the two cases suggests that Kant confused motive and justification. Although the soldier's *motive* in fighting to his death on

14 There remains the puzzle of why Kant commends the suicide of an unjustly convicted innocent man who kills himself to avoid the disgrace of execution (Kant 1979: 155). This is inconsistent with Kant's position on duelling, and with his general ethic of respect. It is further evidence that the genealogy of Kant's ethic of commanding value originates in the honour ethic. It reveals how his applied ethic is a transitional stage from an ethic of honour to an ethic of universal respect.

the battlefield rather than running away may be to protect his honour, he could hardly *justify* engaging in warfare merely to prove his bravery. This is the point of Kant's condemnation of duelling. A just war requires a just cause. States may never initiate war with one another just to prove their honour. A just cause of war can only be the defense of human rights, especially the right to life. Thus, what justifies the soldier's actions is not his concern for his own honour, but his duty to protect life against the enemy's aggression. In a consistent ethics of respect, honour is dependent on respect for humanity rather than the other way around. So neither killing nor letting die is permissible just for the sake of honour. Kant should have applied the same reasoning to the prospective rape victim, who has a right and a duty to protect her own life and thereby defend it against her assailant's aggression. Her concern for her own honour can no more justify fighting to her own death than the soldier's concern for his honour can justify his fighting to his death.[15]

Given these dangers of an ethic that stresses the centrality of moral self-esteem, should we leave behind all concern for it? Shorn of this element, Kantian ethics would turn into contractualism, which upholds universal duties to others but is silent about their application to the self. I do not believe this would be a moral advance. How could it be that others have commanding value to oneself, and one has commanding value to others, but one lacks commanding value with respect to oneself? If one owes duties to all bearers of commanding value, then one has duties to oneself.

Kantian ethics rightly stresses the moral importance of cultivating a sense of personal dignity. If one does not insist on one's own dignity, one can hardly depend on others to pay one respect. This is a claim of social psychology, not of morality. We *must* reject the forfeiture principle (that he who makes himself a worm loses all claim against being stepped on) as a regressive throwback to the ethic of honour. Nevertheless, we must recognize that cultivating a heightened sense of personal dignity – of commanding authority – on the part of the oppressed is critical to securing their human rights. When half of the women surveyed in Zambia allow that husbands have a right to beat their wives for arguing

15 It follows from this that there is no genuine shame to the rape victim in succumbing to rape rather than fighting to her death, since she does what duty requires. Kant's readiness to despise those who submit to degrading treatment for lack of other ways to preserve their lives is further evidence that Kant's ethics grew out of the honour ethic.

with them or burning the dinner, it is no surprise that nearly half of wives surveyed report that their husbands have beaten them.[16] While women must not be held in contempt for failing to stand up for themselves, encouraging them to do so is indispensable. This requires that women cultivate in themselves feelings of their own personal worth as having commanding value, not mere appealing value, and thereby recognize duties to assert their independent dignity. Kant's insight that the distinction between commanding and appealing value is central to morality, and that the former is essentially tied to the authority of persons and is *felt* as well as judged, remains one of his most enduring contributions to moral philosophy.

16 Sharon La Franiere, 2005, "Entrenched Epidemic: Wife-Beatings in Africa", New York Times, 11 August.

The Concept of Love in Kant's Virtue Ethics

Christoph Horn

The moral significance of love is not a prominent topic in Kant's published writings on ethics, although the concept of love – as we know from different sources – formed an important subject matter in Kant's lectures on moral philosophy.[1] This could lead us to the rash conclusion that he treated love in his university lectures as a somewhat exoteric topic, one demanded perhaps by the official curriculum or by the wider public. In the *Doctrine of Virtue*, however, he devotes more than a dozen pages to love [*MM* 6: 401–2 and 448–61] and he adds some considerations on friendship, a relation which he believes to be based both on love and on respect [*MM* 6: 469–73]. These passages have received comparatively little attention from Kant scholars.[2] But since the account of love and friendship in the second part of the *Metaphysics of Morals* can be regarded as a condensed form of what Kant has to say on this topic in the more extended versions of his lectures, it merits additional scruting.

Kant divides, in the *Doctrine of Virtue*, all "duties of virtue" into duties toward human beings and those toward non-human beings, and further classifies the duties toward human beings into those to oneself and those to others [*MM* 6: 413]. Duties to other human beings are again subdivided into "duties of love" and "duties of respect" [*MM* 6: 448]. As is well known, Kant sees moral duties as coercing us to a certain behaviour and not as moral advice which we might follow. This is true

1 Cf., e.g., Herder's lecture notes [27.1: 25–8] and Collins' lecture notes [27.1: 416–30].

2 This situation seems to change in recent literature. See the discussions by Barbara Herman, 1993, *The Practice of Moral Judgment*, Cambridge/London: Harvard University Press, ch. 3; Nancy Sherman, 1997, *Making a Necessity of Virtue. Kant and Aristotle on Virtue*, Cambridge: Cambridge University Press, pp. 224–33; Markus Forkl, 2001, *Kants System der Tugendpflichten. Eine Begleitschrift zu den "Metaphysischen Anfangsgründen der Tugendlehre"*, Frankfurt a. M.: Peter Lang, pp. 201–47; Marcia Baron, 2002, "Love and Respect in Kant's Doctrine of Virtue", in: M. Timmons (ed.), *Kant's Metaphysics of Morals. Interpretative Essays*, Oxford: Oxford University Press, pp. 391–407; and Andrea Esser, 2004, *Eine Ethik für Endliche. Kants Tugendlehre in der Gegenwart*, Stuttgart: Frommann-Holzboog, pp. 374–93.

both for the "duties of right"[3] as well as for the "duties of virtue". So for Kant, there is no difference between the various types of moral duties with regard to the corresponding categorical obligation of the agent.[4] What then are "duties of love"? Can an agent be obliged to practice love? In the second part of the *Metaphysics of Morals*, we are informed that they include obligations of beneficence, gratitude, and sympathy (*Wohlthätigkeit*, *Dankbarkeit*, *Theilnehmung* [*MM* 6: 452]). Furthermore, we learn that love plays an important (if precarious) role in friendship which Kant characterizes as a balanced combination of love and respect: For the person who is on the receiving end of acts of benevolence, love generates a characteristic dependence on the benefactor. The active person, the benefactor, merits gratitude and respect, and for this reason the passive person "stands one step lower than his benefactor" [*MM* 6: 458,15–6]. Since love implies a certain lack of equality, it must be supplemented, in a substantial friendship, by respect [*MM* 6: 470].

So far, the moral importance of love seems to be a very restricted one. But in Kant's work there are a more commonly known group of quotations regarding our topic. In the *Groundwork* as well as in the *Critique of Practical Reason*, we find a famous, if short, exegesis of the biblical commandment (*Mt* 22: 38–9) to love one's neighbor, one's enemy, and to love God. In both texts [*G* 4: 399][5] and [*CPrR* 5: 83], Kant famously rejects the interpretation of the commandment as bidding us to practice "pathological love". What the Bible really demands is, according to both of his texts, "practical love". In the *Groundwork*, we are told that "love as inclination cannot be commanded; but beneficence solely from duty." The adequate form of love is therefore "practical and not pathological love"; it "lies in the will and not in the propensity of feeling, in the principles of action and not in melting sympathy, but the

3 In my opinion, the expression *Rechtspflichten* should be rendered as "duties of right", not as "duties of justice" as they are systematically derived from the rights that one owes to other people. See, e.g., *Collins* [27.1: 415].

4 As Kant puts it in [*MM* 6: 390]: "But a wide duty is not to be taken as permission to make exceptions to the maxim of actions but only as permission to limit one maxim of duty by another (e.g., love of one's neighbor in general by love of one's parents) […]." On this topic, see, e.g., Wolfgang Kersting, 1997, "Das starke Gesetz der Schuldigkeit und das schwächere der Gütigkeit", in: *Recht, Gerechtigkeit und demokratische Tugend*, Frankfurt a. M.: Suhrkamp, pp. 74–120, pp. 112–3.

5 All quotations from the *Groundwork* are taken from the translation by Allen W. Wood: *Kant, Groundwork for the Metaphysics of Morals*, New Haven/London 2002: Yale University Press (hereafter abbreviated as *G*).

former alone can be commanded." Love then must be understood as an attitude motivating us to practice beneficence. Taken in this sense, it is a moral duty which obligates us independently of our given emotional background. Love as a practice of beneficence is not only binding in a situation in which an agent has no positive emotions towards a certain person but also binding if she has a negative feeling towards the person. In the second *Critique*, Kant adds a point lacking in the pertinent passage of the *Groundwork:* he now bases his account of moral virtue as a "progress towards holiness" on this scriptural interpretation. What the Gospel demands through this "law of all laws" is, according to this passage, our approach to moral perfection by an "incessant, but infinite progress" [*CPrR* 5: 83].

Regarding this inhomogeneity of Kant's use of the concept of love, we have good reason to be surprised or irritated: there is, on the one hand, a very restricted sense in which he speaks of duties of love and, on the other hand, a very general sense in which he characterizes love as a summary of all moral laws. What makes things even more complicated is Kant's use of the expression "self-love". There are numerous passages in which he identifies the "principle of self-love" with an inclination in us that is fundamentally opposed to morality, whereas the biblical commandment does clearly include the requirement of self-love.[6]

I. Kant's Concept of Love – Some Basic Features

Thus if we compare the remarks quoted from the *Groundwork* and its parallels in the second *Critique* with what Kant has to say on love in the *Doctrine of Virtue*, a crucial question of interpretation arises: Is the notion of practical love the exact same concept which is dealt with again in the later context? Or does the idea of "duties of love" rest upon another conceptional framework? In all of our three texts love is characterized as beneficence, but the text from 1797 introduces the additional elements of gratitude and sympathy. As difficult as it may be to decide if the meaning of the concept has altered across these three texts, it seems at least obvious that the *Doctrine of Virtue* implies another function of love in morality: and it is, in my opinion, not a broader, but instead a less extended use of

6 An important passage on the immorality of self-love is [*CPrR* 5: 22]. In [5: 73], Kant explains that self-love must be appropriately restricted in order to be in conformity with moral law. Hence, it is not immoral *per se.*

'love'. The reason behind this claim: Whereas Kant interprets the biblical commandment of love – understood in the sense of practical love – as a full equivalent of his own categorical imperative, he speaks of the duties of love as a very particular subset of moral obligations, different from both duties of right and duties of respect (and from several other subsets). What he apparently means by duties of love is a very specific class of obligations an agent has to fulfil. To put it the other way around: According to Kant, the commandment of the Gospel is not restricted to those obligations which he, in his own terminology, calls duties of love.

This observation can be affirmed if we take a look at one of the additional passages in which Kant develops his scriptural exegesis. Perhaps the most positive account of the biblical text is a passage found in *Religion Within the Boundaries of Mere Reason* from 1794, hence roughly in the same period as the *Metaphysics of Morals.* There, Kant describes the ethical doctrine of Jesus Christ as being in fundamental accord with his own moral philosophy. His portrayal of the founder of Christianity includes an interpretation of the commandment of love as a combination of inner and outer obligations [*R* 6: 160,24–161,3]:

> Finally, he [i.e. Jesus Christ] combines all duties (1) in one universal rule (which includes within itself both the inner and the outer moral relations of men), namely: Perform your duty for no motive other than unconditioned esteem for duty itself, i.e., love God (the Legislator of all duties) above all else; and (2) in a particular rule, that, namely, which concerns man's external relation to other men as universal duty: Love every one as yourself, i.e., further his welfare from good-will that is immediate and not derived from motives of self-advantage. These commands are not mere laws of virtue but precepts of holiness which we ought to pursue, and the very pursuit of them is called virtue.[7]

Even if we are not explicitly informed in this passage to what extent Kant considers these scriptural commandments to be congruent with his own position, it is clear that he wishes to identify the moral demands presented by Jesus Christ with those of his own moral system. In this text, we don't find even the slightest objection to love as an inclusive or comprehensive concept, appropriate to describe the entire field of moral obligations. On the contrary, it appears to be a full equivalent of what the moral law requires. Kant says that the commandment to love God possesses the exact meaning of fulfilling one's duties with the right

7 Translated by Greene and Hudson.

motive, and that the demand to love others is identical with acting from a direct, as opposed to an indirect moral reason.

If my interpretation thus far is correct, we now have to face a difficulty which we might call *the problem of the wide or narrow application of love.* How is it possible that Kant uses the same concept to describe, on the one hand, the full range of morality and, on the other, an extremely specific subset of duties? Or is there a difference in meaning between the two uses of love?

Before we can deal with these questions, we have to face another problem when we take a closer look at Kant's remarks on love in the *Groundwork* and the second *Critique.* What should we make of the distinction between pathological and practical love? At least three problems arise with this distinction: [i] The signification of "pathological" is far from being clear.[8] Does Kant allude to the original sense of the term, meaning something like "passive", "passionate" and hence "emotional"? Is pathological love, and only this sort of love, an emotion, as the text of the *Groundwork* suggested? As Kant explains, pathological love is caused in an agent by physical inclination. Does he conclude from this that practical love is something non-emotional?[9] And if the answer is yes, what reason do we have not assume that non-physical (intellectual or spiritual) emotions exist? [ii] For what reason does Kant say that love as inclination, i.e., pathological love, cannot be commanded? What is the difference between the two types of love which makes it possible for practical love to be subject to moral coercion? [iii] Can practical love be identified with beneficence, as the same passage seemed to claim? Apparently a simple identification is excluded. Beneficence is a practice, whereas love is an emotional or non-emotional attitude which someone has (or lacks) towards others (or himself). If this is a correct description of love, then there must be another solution for this difficulty. To sum up our three questions [i-iii]: Does Kant regard love partly as an emotional attitude, partly as a non-emotional one (e.g., as an intellectual or spiritual attitude)? Does he claim that practical love can be commanded, whereas this is excluded for pathological love? And does he believe that the

8 Usually, Kant gives no precise explanation of his use, but renders the expression by something like "natural", "bodily" or "physical".

9 We find a remarkable passage in *On a Newly Arisen Superior Tone in Philosophy* [1795] that goes much further [8: 395–6]. According to this text, an emotion (as well as a pleasure) is pathological if it (genealogically) precedes the moral law, whereas it is appropriate if it is derived from the moral law. The dichotomy pathological – practical, then, goes back to the genealogical source of an emotion.

appropriate form of love is beneficence? At first glance, it looks as if one should answer all of these three questions affirmatively. One may get the impression, both in the *Groundwork* and in the second *Critique*, that he [i] distinguishes between an emotional and a non-emotional sort of love, that he [ii] treats non-emotional love as a possible object of coercion, and that he [iii] identifies non-emotional love with beneficence. But these impressions are false. By way of contrast, in the second part of the *Metaphysics of Morals*, where we find a much more precise account of love, we see that all of these questions must be answered negatively. In order to show this, let me quote the entire passage that begins this treatment of the topic of love (Introduction XIIc: *Love of Human Beings*, [*MM* 6: 401–2]):

> (1) *Love* is a matter of *feeling*, not of willing, and I cannot love because I will to, still less because I *ought* to (I cannot be constrained to love); so a *duty to love* is an absurdity. (2) But *benevolence* (*amor benevolentiae*), as conduct, can be subject to a law of duty. However, unselfish benevolence toward human beings is often (though very inappropriately) also called *love*; people even speak of love which is also a duty for us when it is not a question of another's happiness but of the complete and free surrender of all one's ends to the ends of another (even a supernatural) being. (3) But every duty is necessitation, a constraint, even if this is to be self-constraint in accordance with a law. What is done from constraint, however, is not done from love.
>
> (4) To *do good* to other human beings insofar as we can is a duty, whether one loves them or not; and even if one had to remark sadly that our species, on closer acquaintance, is not particularly lovable, that would not detract from the force of this duty. – But *hatred of them* is always *hateful*, even when it takes the form merely of completely avoiding them (separatist misanthropy), without active hostility toward them. For benevolence always remains a duty, even toward a misanthropist, whom one cannot indeed love but to whom one can still do good.
>
> (5) But to hate vice in human beings is neither a duty nor contrary to duty; it is, rather, a mere feeling of aversion to vice, a feeling neither affected by the will nor affecting it. *Beneficence* is a duty. If someone practices it often and succeeds in realizing his beneficent intention, he eventually comes actually to love the person he has helped. (6) So the saying "you ought to *love* your neighbor as yourself" does not mean that you ought immediately (first) to love him and (afterwards) by means of this love do good to him. It means, rather, *do good* to your fellow human beings, and your beneficence will produce love of them in you (as an aptitude of the inclination to beneficence in general).
>
> (7) Hence only the love that is *delight* (*amor complacentiae*) is direct. But to have a duty to this (which is a pleasure joined immediately to the

> representation of an object's existence), that is, to have to be constrained to take pleasure in something, is a contradiction.[10]

The content of this paragraph might be paraphrased as follows: (1) Since love is an emotional phenomenon, it cannot be something that is at my beck and call; it is not a matter of my willing it into being. Hence, I cannot decide to have it or not to have it, and therefore (given that I have no obligation beyond my capabilities – *ultra posse nemo obligatur*) it cannot be my obligation to have it. (2) There is one sort of love, benevolence, which can be – at least if it is not understood as a feeling – a part of our moral duties. But this kind of love is a moral duty only insofar as what is meant by it is an unselfish attitude; precisely in this sense, however, it is improperly called love. Another widespread, but equally inadequate idea is that love – taken not in the sense of "support of someone's happiness", but in the sense of "complete surrender to another person's goals" – is considered to be a moral duty. (3) Every obligation implies some coercion (if not an outer, then an inner, autonomous coercion). But coercion and love do mutually exclude each other. (4) Good action directed towards others is a moral duty, but that holds true independently of my emotions towards them and regardless of any anthropological facts: even if mankind might not merit being loved, I would still be morally obligated to human beings. Even a misanthropic person (who may not be an attractive object of love at all) must be treated in a fully moral way. (5) To hate the moral vices of other persons is neutral to our obligations and, hence, neither required nor forbidden. It is a mere emotion, and it is neither a consequence nor a cause of an adequate moral attitude. Instead, it is beneficence (*Wohlthun*) that is demanded as our moral obligation. And by practicing it, we begin to love the person who is the object of our beneficence. (6) The biblical commandment "love your neighbor as yourself" does not imply that love is the appropriate starting-point of moral action. It does not even imply that love is an adequate motive at all. It rather describes and demands a process of habituation: If I practice beneficence, I will obtain, as a result, love as a feature of my character. So emotional love can only be a phenomenon of secondary, but not of primary importance for morality. (7) There is one further type of love (besides benevolence), namely,

10 All quotations are drawn from *The Metaphysics of Morals*, translated by M. Gregor; Cambridge 1996. The division in sections is mine.

delight in a person. But since this is based on pleasure, it is not a possible object of our moral duties.

If this paraphrase is correct, one might be tempted to read it as follows: Kant here considers all kinds of love to be pathological, i.e., passive and physical states, and hence as morally inappropriate. But consequently, there would be, for him, no room left for a 'practical' type of love. What he seems to say is that benevolence, insofar as it is a moral duty, is not love, but instead beneficence. Therefore, love is, in each of its forms, morally inadequate. And Kant even seems to relativize or to relinquish the biblical commandment by interpreting it in a very defensive and indirect way. It is surprising to see how little importance he attaches to it. What he seems to provide in this passage, is a strict and complete rejection of the moral importance of love.

Of course, this reading of the text would cause a dramatic dissonance between Kant's later theory of love and the doctrine we found in the earlier passages on pathological and practical love. But fortunately, it can be shown that the reading just proposed is a misinterpretation. Answering our three questions [i-iii], we can see that our later text is in full accord with the earlier texts, but that Kant's entire position is rather different from the one we expected: [i] All types of love are emotional, *amor benevolentiae* as well as *amor complacentiae.*[11] Therefore [ii], no kind of love can be an object of coercion. Morality, however, must be based on coercion. Hence [iii], strictly speaking, love is, in each of its forms, morally inappropriate, since it is an emotion. What is demanded by morality, is not love as an emotion, but instead a practice of beneficence. But since love, understood as benevolence, can arise from a practice of beneficence, we are entitled to say that it *can be* morally appropriate, if only in an indirect manner of *ex post*-habituation. Its *prima facie*-inadequacy is, therefore, not the last word to be said on it. Here lies the good sense of the biblical commandment; what the Bible wants us to realize is that we should first practice beneficence and, thereby, acquire an emotional attitude connected with it or resulting from it. Kant adds that people often misunderstand and exaggerate the significance of love, since they have the wrong idea of what it means in terms of morality to have sympathetic or hostile emotions. For genuine morality, Kant adds,

11 In *Anthropology from a Pragmatic Point of View* Kant distinguishes, more precisely, between affect (*Affekt*) and passion (*Leidenschaft*): emotional love can be rash and honest or intense and blind [*APPV* 7: 252–3. 266. 268. 308].

philanthropic feelings are not required, and misanthropic feelings are not an obstacle.

It is surprising, however, that Kant considers all forms of love to be emotional and hence, at least *prima facie*, morally inappropriate. If I am right about this, then we have found a solution for what I termed the problem of the wide or narrow application of love: both the inclusive and the narrow ways of speaking of love are highly specific ones and do not conflict with each other. The inclusive or comprehensive use of the concept that is typical for Kant's interpretation of the biblical commandment is restricted to the emotional aspect of a possible perfect state of a person. The person under consideration would not only practice beneficence, but would have additionally obtained a (permanent) emotional state of benevolence. In contrast, the narrow or specific sense of love leads to a maxim which is a duty for every agent to adopt: a maxim to practice beneficence, gratitude, and sympathy. This maxim is concerned with some specific moral cases, distinct, for example, from duties of right, duties towards oneself, duties towards non-human animals, etc. Note that, even more so in this second case than in the first, Kant does not make any concessions to love as an emotion.

Some additional light can be shed on the first use of the concept of love if we take a look at a passage found in the *Critique of Practical Reason.*[12] In connection with his interpretation of the biblical commandment, Kant tells us that the Gospel describes "the moral character in its complete perfection" (*die sittliche Gesinnung in ihrer ganzen Vollkommenheit*); but this "ideal of holiness" cannot be attained by any creature. Instead, we are obliged to strive for it in an "infinite progress". The reason why the state of perfection is unattainable: In order to follow all prescripts of morality from love, a creature would have to be completely free from desires and inclinations; acting morally from love would require a nature without physical aspects. Kant then continues [*CPrR* 5: 84,2–21]:

> For, being a creature, and therefore always dependent with respect to what he requires for complete satisfaction, he can never be quite free from desires and inclinations, and as these rest on physical causes, they can never of themselves coincide with the moral law, the sources of which are quite

12 One should note that the passage on the biblical commandment [*CPrR* 5: 83] is immediately preceeded by a refutation of "love for humans" and "love for order" as motivating sources for moral action. It is thus obvious that there is no need to assume a change in Kant's position between 1788 and 1797.

> different; and therefore they make it necessary to found the mental disposition of one's maxims on moral obligation, not on ready inclination, but on respect, which demands obedience to the law, even though one may not like it; not on love, which apprehends no inward reluctance of the will towards the law. Nevertheless, this latter, namely, love of the law (which would then cease to be a command, and then morality, which would have passed subjectively into holiness, would cease to be virtue) must be the constant though unattainable goal of his endeavours. For in the case of what we highly esteem, but yet (on account of the consciousness of our weakness) dread, the increased facility of satisfying it changes the most reverential awe into inclination, and respect into love; at least this would be the perfection of a disposition devoted to the law, if it were possible for a creature to attain it.[13]

What we learn from the text is this: love is a positive and voluntary attitude; it does not fear or apprehend a reluctance of the will towards the law. But since physical creatures are always determined by interests stemming from their physical desires, their happiness cannot be in complete correspondence with morality. Hence the ideal of acting from love and being, thereby, perfectly moral is, at least for us, an illusion; the ideal of holiness is unattainable, given that we have to live under physical conditions. Love is adequate only in one case: if it is not the source, but the consequence of a moral attitude.

Here, however, some subsequent questions arise. When we leave holiness aside, what does it, then, precisely mean to interpret love as a maxim? Why is it impossible that love, taken as emotion, is an object of moral constraint? And for what reason are we obligated to adopt the maxim underlying the "duties of love"?

II. Love Considered as a Maxim

In §25 of the *Doctrine of Virtue*, we are informed that love is morally acceptable if it is taken not as feeling or emotion, but as a maxim [*MM* 6: 449,17–22]:

> In this context, however, love is not to be understood as *feeling*, that is, as a pleasure in the perfection of others; love is not to be understood as a *delight* in them (since others cannot put one under obligation to have feelings). It must rather be thought as the maxim of *benevolence* (practical love), which results in beneficence.

13 Translated by T.K. Abbott.

Is "practical love" a kind of metaphorical glossing-over of beneficence? As we saw in our quotation from the *Doctrine of Virtue*, beneficence cannot be properly called love. Does Kant, for traditional or colloquial reasons, make use of a way of speaking which he usually rejects? No, I think that the key for a correct understanding of what he wishes to say lies in the idea of reformulating love *as a maxim.* But here the difficulties seem to reoccur. How should it be possible to formulate a maxim by which someone imposes love on his or her behavior? A person who adopts a guiding principle such as "If I believe myself to be in pecuniary distress, then I will borrow money and promise to pay it back, although I know this will never happen" [*G* 4: 422], chooses a rule according to which he or she wants to act in a given case and in similar situations in the future. Kantian maxims must be capable of being examined in a universalization test: Either this rule is "merely subjective" (not universalizable), or it is admissible – i.e., either indifferent or an "objective law" (if it is universalizable without contradiction). Now the difficulty is this: On the one hand, the agent cannot choose an emotion since this wouldn't be universalizable. On the other hand, by adopting a maxim of beneficence, he or she would not be doing the same thing as practicing the principle of love. In the second section of the *Groundwork*, the paradigmatically insufficient maxim by which Kant introduces the idea of an imperfect duty towards other people is formulated in this way: "while he [i.e. someone] sees that others have to struggle with great hardships (with which he could help them) – thinks: 'What has it to do with me? Let each be as happy as heaven wills, or as he can make himself, I will not take anything from him or even envy him; only I do not want to contribute to his welfare or to his assistance in distress!'" [*G* 4: 423]. If we extrapolate from this a morally correct maxim which Kant might have in mind, it could read, "If I see others in great difficulties and could help them, then I will contribute to their happiness" (cf. [*MM* 6: 388]). But following Kant's own objection, one should call this a maxim of beneficence, not of love.

What does it, then, mean to say that love must be considered as a maxim and not as an emotion? The central text for the answer to this question is §27 of the *Doctrine of Virtue* [*MM* 6: 450–1]. Kant here characterizes the "maxim of benevolence" as a duty of all human beings to one another and identifies it with the biblical commandment which he calls the "law of perfection". Here, as in his other interpretations of the biblical text, he differentiates between a "practical love of human beings" and the actual possession of love as an emotion (without using the term

"pathological"). After having put aside the supposedly irrelevant question of whether or not certain human beings are "worthy of love", he spells out the sense in which love can have the form of a maxim. In order to correctly understand the moral obligations of human beings towards each other, we should, according to Kant, represent them "by pure reason", namely as a "relation of free actions in accordance with maxims that qualify for a giving of universal law and so cannot be selfish". He continues:

> I want everyone else to practice benevolence (*benevolentiam*) toward me; hence I ought also to be benevolent toward everyone else. But since all *others* with the exception of myself would not be *all*, so that the maxim would not have within it the universality of a law, which is still necessary for imposing obligation, the law making benevolence a duty will include myself, as an object of benevolence, in the command of practical reason.

In this text, we are informed, firstly, what the maxim of love consists in: it is the command to practice benevolence towards all of humanity. The maxim under consideration, we hear secondly, is correctly universalizible and so morally acceptable: we have to imagine the interrelatedness of mankind as determined by mutual benevolence. We are not told why this maxim is morally obligatory, i.e., why it is necessary to adopt it, just as we learn nothing about the questions concerning what it precisely implies and if we should distribute our benevolence equally or by privileging certain persons (I will raise these problems in the next section). However, what we learn, thirdly, is that it includes the person who is examining or testing the maxim. The person examining it must be part of the humanity which is the object of benevolence. Kant next tells us that this does not mean that I am obliged to practice self-love, but instead that it is permitted to include myself in the various objects of my benevolence. This brings us back to a problem raised at the beginning: if self-love is morally intolerable, whereas a self-related form of benevolence is permissible, where is the line of demarcation between these two phenomena?

In order to deal with this question, I would like to turn back to the two best known Kantian texts on love, the passages in the *Groundwork* and in the second *Critique*. Whereas both texts confirm that two types of love exist and that one of them is morally adequate and the other inadequate, they seem to differ considerably in the reasons they give for their interpretation of the Bible. In the *Groundwork*, we are informed that love as inclination cannot be an object of moral coercion. From this, Kant concludes that the sense of the scriptural commandment must be found

elsewhere, namely in practical love. What he means by "practical love" is spelled out, as we have already seen, by "beneficence solely from duty", whereas the meaning of "pathological" is explained by expressions such as "propensity of feeling" and by "melting sympathy". But can Kant really have in mind the idea that it is not possible for emotional love to be the object of a moral obligation? At first glance, we may suppose – as, e.g., Harald Köhl did[14] – that the reason for this lies in the conviction that emotions are not at an agent's disposal, i.e., the agent cannot change them at will. But does Kant really assume that emotions are not under our control, and that they cannot be generated or produced? If this were the case, then we would have to conclude that 'practical love' is something other than an emotion – but shouldn't Kant have claimed that explicitly? Furthermore, can there in fact be a sort of love which is free of any emotional component?

In the second *Critique*, Kant gives us a somewhat different reason for the inadequacy of the first and the adequacy of the second type of love: we are now told that it is impossible to love God through an inclination since He is not a possible object of our senses. Hence, Kant tells us again that the significance of the commandment must lie in practical love. But whereas this conclusion is identical in both texts, two totally different explanations seem to exist for the impossibility to command pathological love.

To make things worse, we thus fail to understand Kant's intention in our two texts. What exactly is the sense of the second explanation? For what reason could we be unable to direct emotions towards a non-sensible object? At the very least, it is by no means obvious why this should be the case. Many adherents of a religion – say Jews, Christians, and Muslims – would vigorously contest the claim that loving God cannot be an emotional behavior. In addition, the first justification of practical love, as we understood it so far, is extremely weak: why should we think of emotions as something purely passive? Emotions are not completely beyond our control. Starting with Aristotle, a long and impressive philosophical tradition exists which cultivated the idea that emotions are not merely passive mental states. According to this tradition, emotions contain an agent's evaluative and situational judgments and can thus be influenced and even transformed. Köhl himself (Köhl 1990, 112–3), in his interpretation of the passage from the *Groundwork*,

14 Harald Köhl, 1990, *Kants Gesinnungsethik*, Berlin/New York: Walter de Gruyter, pp.110–1.

admits that it seems exaggerated to say that emotions simply are not possible objects of agency. An agent has at least some impact on her emotions, and therefore her inappropriate emotions can be, in a given case, the target of serious moral sanctions by her environment. Furthermore, we distinguish natural from unnatural emotions and thus base our moral judgments on this distinction: we expect, e.g., someone to grieve if a person near to her has died.

All of these considerations should make us cautious in providing a rash interpretation of the Kantian dichotomy between pathological and practical love in the *Groundwork* and the *Critique of Practical Reason.* The following short passage to be found in Kant's second interpretation of the biblical commandment is crucial for an adequate reading of this dichotomy [*CPrR* 5: 83,16–22]:

> For a command to like to do a thing is in itself contradictory, because if we already know of ourselves what we are bound to do, and if further we are conscious of liking to do it, a command would be quite needless; and if we do it not willingly, but only out of respect for the law, a command that makes this respect the motive of our maxim would directly counteract the disposition commanded.

Kant's fundamental point is not that emotions cannot be influenced. Like those in the Aristotelian tradition, he accepts – as witnessed by the *Doctrine of Virtue* – a moral education in which emotions are established by habituation. What he regards as absurd is the idea of commanding the generation of feelings which are already present or totally absent. Kant puts the alternative as follows: Either a person already possesses such emotions; then they need not be bidden, but instead restricted, or she doesn't possess them; then they need not be obtained. To command something is, according to Kant, to coerce it; it doesn't make any sense to coerce what is already present in the agent as his natural tendency. Conversely, if I have no emotional love, I don't need to acquire it for moral action. Hence, what is interesting in the quoted text is this: it does not rely on a contrast between moral emotions on the one hand and non-emotional moral motives on the other. Both an inclination (the inadequate motive) and respect (the adequate one) are, for Kant, emotional phenomena. In fact, the contrast in our quotation is based on the distinction between something which we enjoy following as a result of our physical nature, and something which we normally refuse to acknowledge since it stems from morality and is opposed to our physical nature. Moral action cannot be based on our physical inclinations. As is well known, the necessity of a categorical imperative reflects, according to

Kant, an aspect of our anthropological condition, namely, that we do not possess a natural inclination toward what is morally required (cf. [*G* 4: 449 and 454]).

We can conclude that pathological love belongs to the kind of attitude which he usually calls 'the principle of self-love', whereas practical love must be seen as genuine morality. The crucial reason for this is that pathological love is generated by natural inclination; practical love, on the other hand, is derived *ex post* from a firm attitude originally inspired by the moral law. In addition, Kant thinks that emotional inclinations are inappropriate when taken as a basis for moral behavior since they are too unstable to be reliable. But the same holds true for respect: as an emotion it is unstable as well. Emotions are not a possible object of a maxim only insofar as they derive from our physical nature. But that holds equally true for other internal phenomena, e.g., for our interest in happiness and our self-love. In the *Doctrine of Virtue*, we are informed that it would make no sense to coerce someone to strive for his or her happiness [*MM* 6: 386,1–7], and in the same sense it is impossible to oblige someone to practice self-love [*MM* 6: 451,10–2].

III. Why Is It Necessary to Adopt the Maxim of Love?

The main characteristic of maxims which belong to the realm of ethics (in contrast to the realm of right) is, according to Kant, that there is a duty to have them. Whereas non-ethical duties can be voluntarily chosen at first but must then be identified as morally correct or incorrect through a procedure of universalization, ethical maxims are derived strictly from "ends that also are duties". The crucial text for this idea is the following [*MM* 6: 389, 12–26]:

> Only the concept of an *end* that is also a duty, a concept that belongs exclusively to ethics, establishes a law for maxims of actions by subordinating the subjective end (that everyone has) to the objective end (that everyone ought to make his end). The imperative "You ought to make this or that (e.g., the happiness of others) your end" has to do with the matter of choice (an object). Now, no free action is possible unless the agent also intends an end (which is the matter of choice). Hence, if there is an end that is also a duty, the only condition that maxims of actions, as means to ends, must contain is that of qualifying for a possible giving of universal law. On the other hand, the end that is also a duty can make it a law to have such a maxim, although for the maxim itself the mere possibility of agreeing with a giving of universal law is already sufficient.

The standard cases of moral deliberation that Kant describes are procedures of examining a maxim independent of the ends an agent may or may not have. In these cases, the perspective of ends is, according to Kant, irrelevant and even misleading: it is clear from the outset that every human being possesses, on the one hand, an inclination towards his own happiness [*G* 4: 415–6], whereas we have to assume, on the other hand, that no human being has a natural inclination towards the moral law, but instead an inclination towards evil [*R* 6: 36–7]. Since all agents are striving by nature for their happiness, it would be fundamentally mistaken to ground moral philosophy on a "material" (i.e., goal-oriented) principle. All material principles which nature might possibly have given us are derived from the comprehensive end of happiness, in other words from the principle of self-love [*CPrR* 5: 22]. This is the reason Kant gives as to why the categorical imperative in its "Formula of Universal Law" is restricted to a test of universalizability. However, in adopting a correct moral maxim, we modify or even transform our ends. Kant therefore describes, in an important passage of the second *Critique*, the correct procedure of determining one's ends as a "paradox of method in a critique of practical reason". As he explains, "the concept of good and evil must not be determined before the moral law (of which it seems as if it must be the foundation), but only after it and by means of it". And he adds [*CPrR* 5: 63,11–21]:

> Supposing that we wished to begin with the concept of good, in order to deduce from it the laws of the will, then this concept of an object (as a good) would at the same time assign to us this object as the sole determining principle of the will. Now, since this concept had not any practical a priori law for its standard, the criterion of good or evil could not be placed in anything but the agreement of the object with our feeling of pleasure or pain; and the use of reason could only consist in determining in the first place this pleasure or pain in connexion with all the sensations of my existence, and in the second place the means of securing to myself the object of the pleasure.

Concerning ends, there is a necessity to invert the method of moral philosophy, namely first to determine the maxims and, from these, draw conclusions about the ends. This is the reason why the categorical imperative in its "Formula of Humanity" explains the demands of moral law in terms of transformed ends, and not in terms of our natural ends. It demands from us that we change or modify our goals in a way that respects and promotes "humanity" in every human being, including

ourselves.[15] Against this background, we can now see why Kant, in the introduction of the *Doctrine of Virtue*, develops the idea that the doctrine of rights and that of ethics have to take opposite paths [*MM* 6: 382]. The doctrine of rights identifies moral duties in terms of maxims by leaving aside which particular ends someone may have (and is thus confined to a test of universalizability). In contrast, the *Doctrine of Virtue* starts with ends that are, simultaneously, moral duties and derives maxims from them which an agent is obligated to have. We learn that it is a duty for every agent to have one's own perfection and other people's happiness as one's ends. From this, Kant derives the idea of having maxims which are appropriate means towards these ends. Manifestly, the maxim of benevolent love (and that of respect) are, from Kant's perspective, means to achieve the second end: the happiness of others. With this, Kant's justification for the necessity of having the maxim of love becomes clear.

Given that there is a duty to adopt this maxim, we can understand why it is false to reproach Kant for having formulated, in the maxim of love, an egoistic principle – a claim formulated as early as 1839 by Schopenhauer.[16] The duties of love are not derived from an egocentric perspective based on the probability that each of us needs – at least from time to time – the help of other people. They are not, in other words, justified by the well-considered interest of a prudent agent. When taken as a derivation of Kant's second moral end, these duties imply the universal demand to support others on their self-chosen paths towards happiness, since human beings need others to realize the full range of their rational agency.[17] This does not mean that others have the right to demand from us certain, well-defined actions – as in the case of strict obligations. The entitlements which are correlated to our duties of love consist in the fact that others have a right to expect our benevolence. In the same way, we are justly entitled to expect their benevolence.

---

15 I give a more detailed discussion of what may be meant by the "paradox of method" in Christoph Horn, 2002, "Wille, Willensbestimmung, Begehrungsvermögen", in: Otfried Höffe (ed.), *Kant. Kritik der praktischen Vernunft*, Berlin: Akademie, pp. 43–61, at pp. 58–60.

16 *The Two Fundamental Problems of Ethics* [1839/40] 196 – see, on this point, Esser, 2004, pp. 374–6.

17 On this point, cf. the illuminating discussion in chap. 3 of Barbara Herman, 1993, *The Practice of Moral Judgment*, Cambridge/London: Harvard University Press.

At this point, of course, some serious questions must be raised. One of the crucial problems is whether or not we are obligated to support someone in his striving for definite or for arbitrary goals, e.g., whether or not we are obligated to help if he is directed towards self-destructive or amoral goals. Where might an appropriate dividing line between acceptable and unacceptable ends lie? Kant declares that we have the right to refuse to assist someone if we do not believe an end will contribute to her happiness [*MM* 6: 388,5–16], but he does not explain why and in which cases we are allowed to do so.[18] Assuming the acceptability of particular ends, one might go further and ask to what extent and with what intensity we are obliged to assist other people in obtaining their ends. It is neither possible to help everyone nor to help in every situation nor to any extent. Here, Kant's answer simply is that a person in helping should not expend her own resources to the point where she then finds herself in need of help [*MM* 6: 454,2–4]. Admittedly, this approach to a limit based on resources is not really satisfying; we could then ask where the line might be drawn between an adequate and an inadequate form of benevolent love, inadequate insofar as it takes the form of a truly self-sacrificing attitude. One might answer that Kant is thinking primarily of small and daily acts of benevolence, like encouragement, gentleness, patience, giving comfort, empathy and elementary forms of support. Examples might include providing someone with basic information, assisting someone with the lifting or carrying of something, lending someone a tool or something similar, making change for a dollar, etc. But the problem remains as to how to deal with more demanding types of help, such as caring for an ill person or helping someone prepare for his exams.[19] What Kant obviously has in mind with his response to the limit question are cases in which someone needs help in a precarious situation – e.g., a difficult financial position. These cases are rare, and for them the principle not to deprive oneself might suffice.

Furthermore, in what way might an agent integrate duties of love into the general context of all duties? Does Kant think that all strict duties (towards oneself and towards others) have to be fulfilled *before* we have to start taking our duties of love seriously? What should be done in cases of conflict? Kant does formulate the rule that no duty of right can be

18 The problem seems to be that he considers questions of happiness to be thoroughly subjective: e.g. [*CPpR* 5: 36].

19 Kant notably thinks that the parents' obligation to care for their children is a "duty of right", not a "duty of virtue".

disregarded for the fulfilment of a duty of virtue. Further, towards which precise group of persons are duties of virtue directed? Do these duties imply the obligation of everyone to support the ends and interests only of those near and dear to them, thus including perhaps one's neighbors and other persons in one's immediate environment? Is it restricted to a duty to help disoriented strangers who visit your city and incidentally ask *you* for directions? Or does it include the duty to establish a tourist office in *your* city if you can afford that? Is our duty to help thus based on (and restricted to) our range of agency? But if this were true: would it then be our obligation to further this range? And wouldn't it then be a problem that even distant countries are possible areas of our agency – at least given our contemporary technical skills and their long-distance effects? Kant tells us that, insofar as a "love for all human beings" is meant, there is no need to have more of it than the smallest degree, the interest which one should take "is as slight as an interest can be" ("I am only not indifferent […]": [*MM* 6: 451,21–6]).

Even if Kant could provide good answers to all of these questions, one might articulate serious doubts about whether a universal practice of benevolent love would have desirable overall consequences. Does Kant not thereby relinquish the standpoint of moral universality? Love is a selective attitude which privileges some and disadvantages others, and this holds true not only for its emotional, but also for its benevolent form. Kant concedes this point and argues that the possibility of loving in gradations is not excluded since the criterion of the commandment to love one's neighbor is not pathological self-love, but instead, a practical attitude; it is perfectly possible to adopt the maxim of universal benevolence without in all cases fulfilling it to the same degree [*MM* 6: 451–2]. But what about the inequalities and shortcomings of this procedure? It seems likely that, given a universal duty of benevolence, many people would be left without any support from others. Do duties of love therefore include the idea that additional support is needed for those who are not sufficiently helped by others? Are we obliged to establish a network of solicitude in our societies or even throughout the world which would be prepared to counterbalance shortcomings in or lack of individual benevolence?

Kant could respond that these questions are raised from a consequentialist standpoint and that his own moral philosophy is agent-restricted. Love might perhaps not be the right remedy to cure the world; but that is not love's scope. Nevertheless, having a moral duty to love is not left to the agent's own arbitrary judgment; everyone has to

make it their own principle [*CPrR* 5: 83, 6–7]. At this juncture, I can introduce the final point in my interpretation. How should we explain the enormous discrepancies in Kant's evaluation of the moral importance of love? There can be no doubt on the one hand that Kant considers benevolent love to be a complete equivalent of morality (in his accounts of the biblical commandment), and that on the other hand he accuses it of two serious shortcomings: first, love cannot be coerced and its presence is, therefore, contingent (or the result of doing one's duties), and, secondly, beneficent love in a sense humiliates the beloved and in the case of friendship it must be counterbalanced with respect.

The second shortcoming involves a difficulty which I have already mentioned: in the *Doctrine of Virtue*, Kant hesitates on the question of whether love is morally adequate since in a way it makes the recipient dependant on beneficence. He claims that as love implies a sort of lack of equality between two people, it must be compensated by additional respect. Kant views love and respect through the glasses of his philosophy of nature which includes the metaphor of the opposition of attraction and repulsion. Marcia Baron (Baron 2002), however, has asked why such a marked contrast between love and respect is necessary, and it is indeed difficult to see why this should be the case. One might call this *the problem of love and humiliation.* To be sure, Kant's claim would have been more plausible if he had interpreted the relationship between love and respect as a supplementary one: While love demands that I add the happiness of other persons to the rest of my ends, I have to face the problem that I might humiliate them with my support [*MM* 6: 448 below]. Hence, loving others must be balanced by a sound degree of respect. The problem of self-love might be solved in the same way: Whereas love obligates me to practice beneficence, self-respect must save me from sacrificing too many of my own interests [*MM* 6: 450]. Thus, respect is the proper means to correct two major problems of love by balancing love's shortcomings.

IV. A Look at Kant's Lectures on Moral Philosophy

Let us now have a look at an impressive text in which Kant gives us several valuable hints as to why he considers love to be morally insufficient and why it is, nevertheless, an important part of morality. In Collins's notes on Kant's lecture on moral philosophy from 1784/5 we find an extended passage on love [27.1: 413–22]. Its context is again

Kant's systematic account of moral obligations. As before, the issue of love is treated under the heading of 'duties towards other people'. Kant starts by distinguishing between 'duties of good-will or benevolence' and 'duties of indebtedness or rectitude' and discusses the role of love within the former group. As in the case of our first passage from the *Doctrine of Virtue* (quoted above, p. 148 f.), Kant's treatment of love seems to be one-sided and over-simplified, since he generally defines love in terms of inclination: "Love is well-wishing from inclination" [27.1: 413]. Love is benevolence from inclination in contrast to benevolence from obligation which Kant calls beneficence (*Gütigkeit*). All types of love are, according to this passage, based on inclination and stem from an agent's own (physical and egoistic) impulses. Beneficence is not properly love; instead it is based on obligation and derives from 'principles of the understanding'. The person who does something good out of love feels an immediate pleasure about his or her moral action and the well-being of others. In contrast, beneficence implies a mediate pleasure in moral action; in practising it, we experience pleasure from the consciousness of having done our duty. Kant identifies two principal shortcomings of 'well-doing from love' [27.1: 414]. The first is the need to have other people as its objective or target. If I want to help people out of a motivation of love and find no one to help, I will feel dissatisfied and unfulfilled. We can call this love's *lack of autarcy.* The second point is love's discontinuity. If a person who acts from love is deceived, she will be disappointed and may decide to refrain from 'well-doing' in future. We can call this love's *lack of stability.* Both considerations show that it is advantageous to do good works from obligation, i.e., by basing it on principles and not on emotional inclination.

Later on in the same text [27.1: 417], Kant develops a refined scenario of the different forms love can take. He distinguishes between a form of love "that wishes well" and another form "that likes well". This distinction is a complete disjunction; any token of love must belong to one of these two types. Kant then continues by claiming that the "well-wishing love consists in the wish and inclination to promote the happiness of others", whereas the love that likes well "is the pleasure we take in showing approval of another's perfections"; this last form "may be either sensuous or intellectual". This passage sounds as if Kant is now permitting a form of love which is morally appropriate: the love of intellectual liking. However, this impression is false. Kant believes that, like other types of love, this sort of intellectual love that wishes well stems from inclination.

To further clarify his criticism of love as a foundation of morality, Kant undertakes four different thought experiments. He develops them after having claimed that one's "supreme duty is the respect for the rights of others" [27.1: 415]. One gets the impression that what Kant wants to illustrate by contrasting respect and love in this passage is love's radical moral insufficiency. Let us call the four cases: [a] the case of individual negligence of love, [b] the case of universal negligence of love, [c] the case of being individually guided by love, and [d] the case of being universally guided by love.

[a] Suppose, says Kant, that an individual human being would, in his moral agency, act in complete accord with the perspective of right and neglect love altogether [27.1: 415,22–31]. This person "can always close his heart to any other man, and be indifferent to his wretched and pitiable fate". But if this person is "conscientious in observing his bounden duty to everyone", he acts in a completely correct way even if "he gives to no man a jot over his due".

Of course, the person under consideration reminds us of two famous examples in the *Groundwork* [*G* 4: 389]: both the "friend of humanity" (who is temporarily "clouded over with his own grief") and the man who is "by temperament cold and indifferent towards the suffering of others" do their duty in absence of a positive inclination, and their actions have full moral worth. But unlike the *Groundwork*, our text seems even to say that love must be absent if someone's actions should have full moral value. Both passages can be understood as a "morality test": can a moral attitude be perfect in the absence of affirmative feelings and perhaps even in the presence of negative emotions? Kant clearly thinks it can.

[b] Now suppose that all human beings would – in the moral realm – neglect love completely [27.1: 415,31–7]: no one would perform any act of love or kindness. Suppose further that "the rights of every man were left inviolate." Even then "there would be no misery in the world". Kant explains this strong conclusion by adding that he disregards such evil as sicknesses and misfortunes which are not caused by persons; he remarks that "the greatest and commonest of human miseries are due more to men's injustice, than to ill-luck".

One might object that Kant goes too far here, since he fails to add that duties of love must also be established. They are necessary in order to meet the needs resulting not from ill-luck, but simply from the human condition. Be that as it may, we can summarize this consideration as follows: if there were no benevolent love in the world, but instead only a

constant and undeviating sense for the rights of other persons, then an ideal situation would exist – given that natural evils can be disregarded.

[c] In contrast to this last example, Kant asks us to imagine an individual who neglects the rights of others and is guided by love in his moral agency [27.1: 416,16–32]. Kant describes this person as someone who does his duties from benevolence, while disregarding the rights of others. Such a person may, e.g., refuse someone's demand for repayment of a debt owed. In Kant's example, the benevolent agent feels offended by the language of indebtedness, although the creditor is in serious difficulties, and hence refuses to pay. Kant concludes that this singular moral failure is worse than "all the kindly and beneficent acts he has done over a lifetime […] for this is a wholly different sort of reckoning, in which those acts have no place at all".

In this passage, our impression is confirmed that Kant does not only regard benevolent love as having no moral worth, but that he considers it to be a possible obstacle to real morality. Someone who practices benevolence will easily feel hurt if his love is not appropriately acknowledged or mirrored. This feature of the loving person shows, according to Kant, that benevolence is normally based on self-love.

[d] Finally, Kant asks us to imagine a situation in which all humans are guided exclusively by love [27.1: 416,32–417,8]. In this case, Kant contends, there would be no private property at all (no "mine and thine"), and the world "would be a stage, not of reason, but of inclination". Kant's point is that in such a world no one would have a sufficient motive for working, but instead everyone "would rely on the charity of others". Even if "the greatest abundance of everything" existed, this situation would be intolerable, he claims, since people would be in a merely passive state; they would – akin to children – enjoy the various goods for as long as they are available.

Again, this passage leads to a serious objection. Kant here connects two completely different things: the moral problem of adequate and inadequate motives for action and the prudential question of which motives are best for the fulfilment of human (individual and social) interests.

If my comments on these texts are correct, these four thought experiments are far from satisfying, but at least we find in them a strong support for our interpretation. The four cases under consideration provide an extremely critical account of what is going on between the standpoint of right and the standpoint of love. To put it in a somewhat pointed way, [a] Kant pleads for a rigid form of individual moral

correctness; [b] he neglects the needs stemming from human nature, marginalizes socioeconomic problems, and perhaps underestimates the importance of physical evil; [c] he makes use of a stereotype in criticizing an imaginary moral enthusiast for his alleged disobedience to morality; and finally, [d] he overstates the importance of possessive individualism by confusing questions of efficacy with the demands of morality. After taking a closer look, however, it seems clear that Kant introduces these four cases to sharply contrast an adequate and an inadequate form of moral motivation. Must we thus draw the conclusion that Kant rejects the moral significance of love in a radical way? In a sense, yes, but there is one further thing to be said. Up until now, in my treatment of Collins's lecture notes from 1784–5 I have left out a remarkable passage. In it Kant describes our inclination to love and benefit others from some teleological or providential point of view and maintains that we all take part in some form of "general injustice" (*allgemeine Ungerechtigkeit*) [27.1: 415,37–416,16]:

> But since respect for rights is a result of principles, whereas men are deficient in principles, providence has implanted in us another source, namely the instinct of benevolence, whereby we make reparation for what we have unjustly obtained. We thus have an instinct for benevolence, but not for justice. By this impulse, men take pity on another, and render back the benefits they have previously snatched away, though they are not aware of any injustice; the reason being that they do not rightly examine the matter. One may take a share in the general injustice, even though one does nobody any wrong by civil laws and practices. So if we now do a kindness to an unfortunate, we have not made a free gift to him, but repaid him what we were helping to take away through general injustice. For if none might appropriate more of this world's goods than his neighbour, there would be no rich folk, but also no poor. Thus even acts of kindness are acts of duty and indebtedness, arising from the rights of others.[20]

Two aspects of this text seem to be especially interesting and merit our close attention: Firstly, Kant develops a kind of teleological justification of love (in the sense of benevolence)[21] that tries to uncover the hidden meaning of our natural inclination to philanthropy: benevolent love is not a dysfunctional or an amoral attitude as it may seem at first glance. On the contrary, it performs a respectable moral function. Secondly,

20 Translated by Peter Heath.

21 Even erotic love can be, for Kant, part of a teleological story that leads to respect and morality. In the *Speculative Beginnings of Human History* [8: 112–3] we find a short passage in which Kant discusses the influence of reason on our erotic and sexual behavior.

benevolent love is nothing but a compensation for a general injustice, a state in which we all are involved. By supporting someone who is in trouble, we do no more than to reimburse this person for his or her losses under the present conditions of general injustice. Two pages earlier, on 413–4, we are told why Kant regards the distribution of goods in the social world as a situation of far-reaching injustice. Human beings are, according to this text, put into the natural world as its "guests"; concerning the goods we find in nature everybody has the *right* to enjoy them. The explicit terminology of rights underlines that what is at stake here is a moral entitlement. And Kant expressly adds that everyone has an *equal* right to the goods of this world. But since "God has not parcelled out his share to anyone, but has left it to men to divide these goods among themselves, everyone must so enjoy these good things to live, that he is mindful also of the happiness of others […]" [27.1: 414]. Kant illustrates this point by the imaginary situation that someone finds a "table laden with food in the forest." He claims that in this situation the finder is not entitled to eat the entire meal. Hence, my own partaking of natural goods must be restricted by the idea that others have an equal right to participate. Kant concludes that "I have obligations to limit my consumption, and to bear in mind that nature has made these arrangements for everyone. This is the source of well-doing by reason of obligation."

It is no exaggeration to say that our text contains a Kantian argument for an "obligation to solidarity". Nevertheless, one shouldn't make too much of it; it contains no support for the claim that Kant defends the idea of a welfare state.[22] But in light of this passage, we can at least relativize the impression that Kant marginalizes poverty and is uninterested in socioeconomic questions. This should be revised in the following way: Given that every human being has, on Kant's view, an equal right to benefit from the goods of the world, we have a moral obligation to respect the genuine interests of others pertaining to their

22 The thesis that Kant, to a certain degree, supports the idea of a welfare state has been defended by J. C. Merle (1999, Funktionen, Befugnisse und Zwecke der Staatsverwaltung, in: Otfried Höffe (ed.), *Metaphysische Anfangsgründe der Rechtslehre*, Berlin: Akademie, pp. 195–212, at 203–6) and by K. Steigleder (2002, *Kants Moralphilosophie. Die Selbstbezüglichkeit reiner praktischer Vernunft*, Stuttgart/Weimar: J.B. Metzler, pp. 215–22) against the majority of Kantian scholars. But our passage provides no evidences for this interpretation, since Kant doesn't speak of (institutional or individual) obligations which correspond to rights to an equal share.

attainment of well-being and happiness.[23] Now, if we have moral obligations to benefit others until they possess a fair portion of the goods in the world, there must be a sense in which we can fulfil our duty. It must be possible that we oppose our involvement in general injustice in a morally adequate way and not only by being outsmarted by nature, who – to her credit – has implanted in us some philanthropic inclinations. This is the reason why Kant demands, in the same text, that "all moralists and teachers should […] represent acts of benevolence to be acts of obligation, and reduce them to a matter of right" (417). That this demand is not cynical in nature is clear when one considers the passage in which Kant defended the idea of everyone's equal right to the goods of the world. He certainly does not want instructors to *lie* in interpreting actions of benevolence in terms of duties.

There is one additional aspect which further increases the value of Collins's lecture notes for our topic: On a close reading, Kant does not just reformulate our inclination to benevolent love in terms of an obligation to solidarity. He instead goes further and leaves room for love as a moral emotion. In 417, after having claimed that well-wishing can be commanded only as a duty (and not as love), he describes a way of attaining adequate emotions. In a sort of short moral pedagogy, Kant proposes that we should perform acts of benevolence in the consciousness of being obliged, and that we can, by such a practice, acquire a sort of moral habituation. The passage reads as follows [27.1: 417,11–9]:

> If, however, we do well by someone from duty, we get used to this, so that we subsequently do it from love and inclination as well. If we speak well of someone, simply because we see that he deserves it, we get used to this, so that we afterwards intone his merits in everything. Thus even love from inclination is a moral virtue, and might be commanded to this extent, that one should first practice well-doing as a duty, and later, through habituation, out of inclination as well.

We encountered the idea of a habituation which establishes moral emotions in an agent in our first quotation from the *Doctrine of Virtue* (above, p. 148 f., point (6)). In the above quotation, this point is strongly confirmed: there is a way in which emotional love can be characterized as

23 Kant does not exactly repeat this line of thought in the *Doctrine of Virtue*, but a passage exists which is quite close to it [*MM* 6: 458,1–11]. Here, too, he raises the question as to what the world would be like if there were only duties of right and they were observed conscientiously. His answer is that benevolence, which is "a great moral adornment", would then "be missing in the world".

morally adequate even if it is in a secondary and indirect way. Someone who practices benevolent love towards other people (love that wishes well), can become used to thinking and speaking well of other people. According to Kant, specific moral feelings exist which, like other emotions, serve as indicators of pleasure and displeasure, but it is characteristic for them that they are a "subjective effect" of a moral determination of the will.[24] He calls this sort of habituation explicitly moral virtue. As is well known, especially from the more extensive discussion in the *Critique of Practical Reason* [5: 82–6], Kant regards this process of habituation – and not its result – as moral virtue. Virtue therefore is not identical with holiness; instead, it means "strength to overcome obstacles" and is in a certain sense, following Kant's thoughts to their completion, even more remarkable than actual moral perfection [*MM* 6: 396–7].

Kant, thus, admits only an indirect or secondary role for love and other moral feelings. This observation can be corroborated by a final observation. In the *Doctrine of Virtue*, Kant distinguishes between two kinds of sympathetic feelings which he terms *humanitas practica* and *humanitas aesthetica* [*MM* 6: 456–7]. Whereas the first signifies the capacity to communicate one's feelings to another person, the second signifies one's receptivity for the emotions of another person. Kant accepts the first sort by reminding us of the Stoic ideal of the sage.[25] This capacity lends us the possibility of leading the life of a moral hero. He rejects, however, the second capacity by raising the objection that such a feeling needlessly multiplies the evil in the world. Being morally irritable or receptive possesses no intrinsic worth for Kant. All moral emotions that have value must be derived from practical reason.[26]

24 See [*MM* 6: 221]: "[…] there arises the concept of a duty, observance or transgression of which is indeed connected with a pleasure or displeasure of a distinctive kind (a moral *feeling*), although in practical laws of reason we take no account of these feelings (since they have nothing to do with the *basis* of practical laws but only with the subjective effect in the mind […]."

25 On this topic see Marcia Baron, 1995, "Sympathy and Coldness: Kant on the Stoic and the Sage", in: *Proceedings of the VIIIth International Kant Congress*, vol. 1, part 2, Milwaukee, pp. 691–702.

26 I would like to thank Ryan Bremner for numerous corrections and improvements of my English text.

Kant on Respect, Dignity, and the Duty of Respect

Stephen Darwall

> Any respect for a person is properly only respect for the law … of which he gives us an example… [W]e represent … [him] as, so to speak, an *example of the law* (to become like him in this practice), and this is what constitutes our respect [*G* 4: 40ln].[1]
>
> *Respect* is always directed to persons, never to things… [B]efore a humble common man in whom I perceive uprightness of character in a higher degree than I am aware of in myself *my spirit bows…* His example holds before me a law that strikes down my self-conceit… *Respect* is a *tribute* that we cannot refuse to pay to merit, whether we want to or not; we may indeed withhold it outwardly but still we cannot help feeling it inwardly [*CPrR* 5: 76–77].[2]
>
> But a human being regarded as a *person*, that is, as the subject of a morally practical reason, is exalted above any price … he possesses a dignity (absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. He can measure himself with every other being of this kind and value himself on a footing of equality with them.
>
> Humanity in his own person is the object of the respect which he can demand from every other human being, but which he must also not forfeit [*MM* 6: 434–435].[3]

Here is a familiar story, one I've told often myself. Kant bequeathed to modern moral thought the doctrine that all rational beings or persons have a dignity that gives them an equal claim to a respect that differs from any we accord to any form of merit, even moral merit. Frequently this characterization is put forward on the basis on Kant's most familiar ethical writings, *Groundwork* and *The Critique of Practical Reason.* When, however, one looks carefully at these works, a much more complicated

---

1 Immanuel Kant, 1996, *Groundwork of the Metaphysics of Morals*, in: Mary Gregor, trans. and ed., *Practical Philosophy*, Cambridge: Cambridge University Press. Further references will be to volume and page number of the canonical *Preussische Akademie* edition and placed parenthetically in the text (hereafter abbreviated as *G*).

2 Kant, 1996, *Critique of Practical Reason*, in *Practical Philosophy* (hereafter abbreviated as *CPrR*).

3 Kant, 1996, *Metaphysical First Principles of the Doctrine of Virtue* (*Metaphysics of Morals*), in *Practical Philosophy* (hereafter abbreviated as *MM*).

and puzzling picture emerges.[4] As often as not, Kant characterizes the dignity of persons as a species of merit rather than a standing that persons have regardless of merit.

For example, in *Groundwork* II, just after the famous passage in which he says that "morality, and humanity insofar as it is capable of morality, is that which alone has dignity," Kant adds that it is the "practice" of following the moral law that is the "object of an immediate respect" and that "this estimation lets the worth of such a cast of mind be cognized as dignity" [*G* 4: 435]. Kant does also say, of course, that "rational 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)" [*G* 4: 428]. But he follows this less than ten pages later by saying that it is a "morally good disposition" that makes a rational being "fit to be a member of a possible kingdom of ends" [*G* 4: 435] and that what distinguishes "rational nature" from the rest of nature is that it sets itself "an end" that is "the matter of every good will" [*G* 4: 437]. It is clear from the context that in this latter passage he is not saying, as he is often interpreted, that what makes rational beings ends in themselves is their having the capacity to set ends in general, but that they actually set themselves the specific end of respecting the moral law, that is, the end that defines the good will.[5] Taken together, these passages can encourage the thought that persons are ends in themselves only to the extent that they actually follow the moral law, giving a different resonance to the familiar passage, quoted above, that humanity has dignity "only *insofar as* it is capable of morality" ([*G* 4: 428], emphasis added).

My aim in what follows is, first, to analyze what Kant says about respect and dignity in the *Groundwork* and *The Critique of Practical*

4 I was initially set thinking along these lines by conversations with Oliver Sensen. See, especially, Oliver Sensen, 2004, "Kants Begriff der Menschenwürde" in: F.-J. Bormann/ C. Schröer (eds.), *Abwägende Vernunft*, Berlin: De Gruyter, pp. 220–236; and Sensen, 2004, "Freedom and the Categorical Imperative: Kant's Conception of Human Dignity," Ph. D. Diss, Cambridge University, which brings out some of the complexities, along with a very interesting account of the history of the conception of dignity by which Kant was himself influenced. See also Stephen Massey, 1983, "Kant on Self-Respect", in: *Journal of the History of Philosophy*, vol. 21, pp. 57–74.

5 See, for example, Christine Korsgaard, 1996, "Kant's Formula of Humanity", in: *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, pp. 114–132. For this point see Arthur Melnick, 2002, "Kant's Formulations of the Categorical Imperative", in: *Kant-Studien*, vol. 93, pp. 291–308.

Reason in order to bring out how complex and divergent from the familiar story his remarks there are. This will set the stage for the pursuit of my second aim, the examination of a substantially less familiar aspect of Kant's writings on respect and equal dignity, his discussion of the duty of respect in the *Tugendlehre* in *The Metaphysics of Morals*. I shall argue that there we do find pretty unambiguously the idea that all persons have a dignity that warrants equal respect regardless of merit. Even more strikingly, we find the thesis that we all have an equal standing to *demand* respect from one another. I shall argue that this claim goes significantly beyond anything that can be found in the *Groundwork* or the second *Critique*, but that it is also one whose full significance can be read back into these earlier works in a philosophically interesting way. Even here we shall find some residue of the earlier passages that connect respect for the special worth of persons to the response to merit, however, since Kant says that the respect we have a duty to accord all persons is one the vicious person is not really "worthy of" [*MM* 6: 465]. This might suggest that fulfilling the duty involves dissimulation, treating the vicious person as though she were worthy of an attitude she is not. It will turn out, however, that Kant cannot really mean this. Since the duty of respect is an *ethical* rather than a legal or juridical duty, it can only be satisfied by something internal, not by any form of action toward others regardless of our subjective state. But, by the same token, in order to be the object of a duty, respect must be something we can voluntarily adopt, and therefore not a feeling or a spontaneous attitude like the response to merit. The duty of respect is a duty to have a specific *maxim*, the maxim of treating all persons as ends in themselves regardless of their merit, thereby according one another the authority to make demands of each other as equal persons.

I. Dignity and Respect in the Groundwork

Respect's first appearance in the *Groundwork* is in Chapter I in the course of an elaboration of Kant's famous claim that only actions that are done "from duty" have moral worth. When an action is done from duty, Kant says, what determines the will "objectively" is simply the moral law itself; but "subjectively" the will is determined by "*pure respect* for this practical law" [*G* 4: 400]. Kant is here distinguishing between what a morally good agent takes as a sufficient ground or reason for her action from a practical point of view, on the one hand, and what operates as a cause of

her behavior within her empirical psychology, on the other. And he anticipates the specific account of the role of the *feeling* of respect in the "phenomenal" dimension of his transcendental practical philosophy that he will provide in *The Critique of Practical Reason.*[6] We will have occasion to examine this in some detail below, but it will be helpful to have its broad outlines before us now. Genuinely moral action requires transcendental freedom, that the will is a *causa noumen* itself [*CPrR* 5: 55]. Noumenal *practical* causation is free choice, and so, although any causation involves a law of some kind, free choice must involve determination by a *practical law.* It is the moral law itself, then, that grounds free moral action; that is what determines a good and morally worthy willing "objectively." But insofar as it can be an object of experience and part of the empirical world, this choice and action must also have phenomenal aspects, and these phenomenal aspects must have phenomenal causes. Kant holds that the phenomenal cause of phenomenal moral choice and chosen action is the feeling of respect for the moral law. Later he will call this phenomenal feeling of respect "*reverentia*," [*MM* 6: 402] and in the second *Critique*, he provides a fascinating account of its empirical psychology.[7] Here though he simply states that this feeling determines the morally good will "subjectively," and we should understand this claim in terms of this later account. Finally, in addition to the subjective and objective determiners of moral action, there is the action's *maxim*, its "subjective principle of volition" [*G* 4: 40ln]. This, Kant tells us, is "complying with such a law even if it infringes upon all my inclinations" [*G* 4: 401]. So moral action invariably involves the following three elements: the action's "objective" determiner – the moral agent's reason for acting which is, quite simply, the fact that it is required by duty or the moral law; its "subjective" determiner – the feeling of respect for this law; and its maxim, which

6 On this contrast, see Andrews Reath, 1989, "Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination", in: *Kant-Studien*, vol. 80, pp. 284–302; and Patrick Frierson, 2005, "Kant's Empirical Account of Human Action", in: *Philosophers' Imprint*, 5, <http://www.philosophersimprint.org>.

7 As I note in the next paragraph, Kant contrasts *reverentia* with *observantia* or practical respect, a form of respect that is manifested in the agent's *maxim.* As we shall also see, the duty of respect concerns *observantia* in the first instance, although, as with any moral duty, acting *from* this duty, or from the moral law that includes it, invariably involves *reverentia*, the feeling of respect, at the phenomenal level.

encodes the agent's reason – complying with, or to comply with, this law for its own sake regardless of inclination costs.

As will become clearer in the second *Critique*, the feeling of respect (*reverentia)* for the moral law is the phenomenal or felt aspect of respect for law, considered practically – of *observantia* or "respect in the practical sense," as Kant calls it in *The Metaphysics of Morals* [*MM* 6: 449]. Respect for law, as it features in the morally good agent's *maxim*, is not a feeling. An agent respects the law in this sense, when she sets out to comply with it for its own sake, that is, when she treats the fact that an action is required by the law as sufficient reason to comply with it, regardless of inclination costs, and complies with the law for this reason. This practical or *observantia* form of respect for law is realized, not in any feeling, but in the agent's own choice and the practical reasoning on which it is based. Moral action thus invariably involves both a feeling of respect and the practical respect for the moral law (*observantia*) of which it is a feeling.[8]

Practical (*observantia*) respect for the moral law is thus an instance of the kind of respect that I have elsewhere called *recognition respect.*[9] Unlike respect that responds to merit, which I there called *appraisal respect*, recognition respect is realized in recognizing its object's authority and therefore shows itself in how we regulate our conduct toward the object. We respect the law in this sense, when we recognize its authority in deliberating about how to act in relation to it, by, as Kant supposes, treating the moral law as sufficient reason to act regardless of our inclinations. Recognition respect is thus a distinctively *practical* form of respect (or *observantia*) and is different therefore from any spontaneous response to merit (appraisal respect) or even from the conviction that something, say the moral law, warrants recognition respect.

Consider in this light now what Kant says at this juncture about respect for persons in the passage at the beginning of this essay.

> Any respect for a person is properly only respect for the law … of which he gives us an example… [W]e represent … [him] as, so to speak, an *example of the law* (to become like him in this practice), and this is what constitutes our respect [*G* 4: 40ln].

Kant has not yet introduced the ideas of humanity as an end-in-itself or the equal dignity of persons. At this point in the *Groundwork*, he is

8 See, e.g., [*MM* 6: 464].

9 Darwall, 1977, "Two Kinds of Respect", in: *Ethics*, vol. 88, pp. 36–49.

concerned with a distinctive kind of evaluative response, moral esteem, as I read him, that an "impartial rational spectator" would have toward a good will and actions that are done from duty (thus that manifest respect for the moral law) [*G* 4: 393]. One might therefore expect to find him saying that this esteem can itself be seen as respect, but of a different kind – (appraisal) respect for moral merit or moral esteem – albeit a kind that is itself grounded in recognition respect for the authority of the moral law. But this is not what he says. For reasons that aren't entirely clear, he prefers to reserve 'respect' at this point for (recognition) respect for the law. His picture seems to be the following. When we feel moral esteem for a person of good will, say in light of an action of hers that has moral worth, we see this person as an example of the law. She represents, we might say, both what the law requires and the way in which it requires it, since her maxim is to comply with the law, countervailing inclinations to the contrary notwithstanding. We consequently understand the law to require us to be like her (we might add: countervailing inclinations to the contrary notwithstanding). In so orienting ourselves toward her we in effect make respect for law our own maxim and, in so doing, we begin to realize (practical) respect for law ourselves. Again, as will become clearer in the second *Critique*, this practical respect must have a felt phenomenal aspect, which Kant here refers to as a "feeling" that is "*self-wrought* by means of a rational concept" [*G* 4: 401n]. But the respect of which it is a feeling is practical respect for law.[10]

So far as I can see, Kant's semantic preference at [*G* 4: 401n] to reserve 'respect' for recognition for the moral law's authority is peculiar to this passage. In the second *Critique* he writes that "*respect* is a *tribute* that we cannot refuse to pay to merit" [*CPrR* 5: 77]. Presently we shall consider a number of passages where Kant seems to link (and in some cases to identify) the dignity of persons with moral merit. But we should just note first that, again, so far as I can see, nothing of philosophical substance would be lost if he were simply to have said that there is a form of moral esteem that is also often called "respect," albeit of a distinct (merit-responding, appraisal) kind and that this is nonetheless parasitic on (recognition) respect for law. What we esteem (respect) in the person of good will is her respect for the moral law. This is what her merit consists in.

10 It is worth noting that there presumably must be both noumenal and phenomenal counterparts of the esteem (later, respect) that responds to merit also.

I turn now to what Kant says about the dignity of persons in the *Groundwork.* Kant introduces the topic with his famous contrast between "price" and "dignity" [*G* 4: 434]. Whatever has a price "can be replaced by something else as its equivalent." What has dignity "is raised above all price" and "admits of no equivalent." Things that satisfy "human needs and inclinations" have a "market price." That which brings delight, without even any prior need or inclination has a "fancy price" [*G* 4: 434]. Contrasting with these things of "relative worth," is "that which constitutes the condition under which alone something can be an end in itself." This has an "inner worth, that is, *dignity*" [*G* 4: 435]. There follows the familiar passage, part of which we quoted at the outset.

"Now, morality is the condition under which alone a rational being can be an end in itself, since only through this is it possible to be a lawgiving member in the kingdom of ends. Hence morality, and humanity insofar as it is capable of morality, is that which along has a dignity" [*G* 4: 435].

It is common here to stress "*capable* of morality" and to interpret Kant as saying that the distinctive worth or dignity he is ascribing to rational persons is one they have regardless of their actual conduct or maxims, hence regardless of their merit. But consider what Kant then says. First, he contrasts skill and diligence, and lively imagination and humor, all of which have their price (the first two a market price, the second two a fancy price) with "fidelity in promises and benevolence from basic principles (not from instinct)," which "have an inner worth" or dignity [*G* 4: 435]. The worth of the latter, he says, consists not in "effects arising from them," but in "maxims of the will that in this way are ready to manifest themselves through actions" [*G* 4: 435]. The contrast here is clearly intended to recall Kant's claims about the good will and morally worthy actions, whose distinctive worth lies "nowhere else *than in the principle of the will*," regardless of its effects [*G* 4: 400].

Kant follows this with the passage I quoted at the outset. "[T]he will that practices" such (morally worthy) action, Kant says, is "an immediate object of respect" [*G* 4: 435]. And "[t]his estimation" (presumably that involved in such moral esteem or respect), he continues, "lets the worth of *such a cast of mind* be cognized as dignity and puts it infinitely above all price" ([*G* 4: 435], emphasis added). And if this isn't enough to suggest strongly that Kant is, at least at this point, thinking of the "inner worth" that he here calls dignity as the same as the "unconditional inner worth" that he claims that good will or the disposition to morally worthy action uniquely have, at the beginning of the *Groundwork* [*G* 4: 393–394],

Kant begins the next paragraph with the question: "[W]hat … justifies a morally good disposition, or virtue, in making such high claims?" [*G* 4: 435]. His answer is that "it is nothing less than the *share* it affords a rational being *in the giving of universal laws*, by which it makes him fit to be a member of a possible kingdom of ends" [*G* 4: 435].[11]

These are remarkable passages; at least, they are when they are taken by themselves and read against the background of the familiar interpretation with which I began. First, in contrast also with the beginning passages we considered according to which respect for persons is never respect for their moral merit or meritorious action, these passages say that respect for the dignity of persons is, indeed, respect for their merit, or, at least, for the merit of which they are capable. This conflicts with what many, myself included, have taken to be the entire tenor of the Kantian view, that the capacity for moral action which any person, virtuous or vicious, gives persons a dignity that is no form of merit, and respect for which is of a fundamentally different kind, recognition respect, than the appraisal respect that merit or virtue inspires to the degree of its merit or virtue. Second, Kant here clearly seems to say that being an end in itself and thereby fit to be a member of the kingdom of ends is conditional on having morally worthy maxims, that is, on having "virtue" or "a morally good disposition."

We should therefore next consider what Kant says when he introduces the idea of persons as ends in themselves in connection with the Formula of Humanity (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.") [*G* 4: 429]. Here we find the familiar doctrines that human beings and "in general every rational being *exists* as an end in itself" and that "rational beings are called *persons* because their nature already marks them out as an end in itself…, and hence so far limits choice (and is an object of respect)" [*G* 4: 428]. These passages certainly encourage the common interpretation with which I began, namely, that rational persons are appropriate objects of respect irrespectively of their merit, hence, that they are fitting objects of a kind of respect that differs from that which responds to merit. And there is nothing in the passages in which Kant defends, illustrates, and elaborates FH to put one off this interpretation. Kant's "derivation" of FH, much discussed by recent commentators, his application of FH to

11 Also: "[W]e … represent a certain sublimity and *dignity* in the person who fulfills all his duties" [*G* 4: 440].

examples, and his remarks just following all seem to encourage an interpretation within the usual range [*G* 4: 427–431].[12] Our ends can furnish categorical rather than hypothetical imperatives only by virtue of their relation, not "to a specially constituted faculty of desire," but to rational nature itself [*G* 4: 428]. Everyone necessarily "represents his own existence" as a rational being as an end in itself [*G* 4: 429]. The persons toward whom we act must be able to share "the end" and "principle" of "the very same action" [*G* 4: 430]. And so on.

Despite all this, we then find Kant saying a few pages further on that "morality is the condition under which alone a rational being can be an end in itself," and that it is a "morally good disposition" that "makes [a rational being] fit to be a member of a kingdom of ends" [*G* 4: 435]. And then there are the passages identifying dignity with moral worth (merit) along with the passage I mentioned briefly near the beginning that what distinguishes "rational nature … from the rest of nature" and gives rational beings a "dignity … over all merely rational beings" is "that it sets itself an end" that is "the matter of every good will" [*G* 4: 437–438]. My purpose in pointing to these passages is not to propose a revisionist interpretation of FH, nor even to recommend that we look for one.[13] Neither do I mean to suggest that there is no way of squaring the passages with familiar interpretation of FH and the equal dignity of rational persons. My purpose rather has been simply to show that this is more difficult than is usually believed. What Kant actually says in the *Groundwork* about equal dignity and respect is substantially more complex and, in some ways, strange, than the familiar picture would have it.

II. Respect in the Critique of Practical Reason

In Chapter III ("Of the incentives of pure practical reason") of Book I of the second *Critique*, Kant provides an empirical psychology of moral action, an account of how moral choice and conduct can be causally

12 See, e.g., Thomas E. Hill, Jr., 1991, "Humanity as an End in Itself", in: *Ethics*, vol. 80, pp. 84–99; Christine Korsgaard, 1996, "Kant's Formula of Humanity", in: *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press; and Allen Wood, 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press, pp. 111–155.

13 It would be interesting, however, to consider whether there might be one that fits with the passages I have mentioned when they are taken at face value.

determined in the phenomenal realm even though, considered noumenally, it is transcendentally free. The central philosophical idea, which I mentioned briefly earlier, is Kant's "two standpoints" doctrine that a free moral action, considered noumenally, manifests itself phenomenally as having a phenomenal cause. From the practical point of view, the ground of moral action, the agent's reason, is simply the moral law itself. And the agent's maxim is "to comply with (that is, to respect) the moral law." But the manifestation of moral action in the phenomenal realm must have a cause in the agent's empirical psychology, which Kant supposes must be a feeling. Kant concludes that it is an *a priori* truth that there must be a feeling that plays this functional role, and in Chapter III he shows how respect can do so. Since "respect for the moral law is a feeling that is produced by an intellectual ground," it is a feeling "that we can cognize completely *a priori*" [*CPrR* 5: 73].

Chapter III proceeds to lay out a fascinating phenomenology of the feeling of respect, the overall aim of which is to show how, although this feeling is the phenomenal cause of moral action, it is not itself "the incentive to morality" (i.e., what the moral agent herself takes as a reason to be moral). The feeling of respect is rather "morality itself subjectively considered as an incentive" [*CPrR* 5: 76]. The key psychic insight in Kant's phenomenology of respect turns on a distinction between self-love and self-conceit and on an account of how the feeling of respect consists in feeling the moral law's capacity to "humiliate" self-conceit and "awaken" respect for itself (and for its source in pure practical reason) [*CPrR* 5: 74–75].

Self-love, in this context, is the "natural propensity" to take merely "subjective determining grounds" of the will to have objective normative significance. Like a naïve experiencer who takes an apparently bent stick in water to be really bent, a naïve agent may take his desire's object to be a source of reasons, oblivious to peculiarities of the perspective that his desire gives him. Things that are important to him seem (to him) to be important period. Self-love in this sense need pose no deep threat to morality; in principle, it need be no more dangerous than the innocent mistakes of perspective that are correctible once we can draw a subjective/objective distinction within our experiences and accept some as mere appearances. Self-love, Kant thinks, needs only to be curbed by the moral law.

Self-conceit, on the other hand, assaults the moral law directly, and so it must be "humiliated." It is a form of arrogance (*arrogantia*): the presumption that one has a kind of worth or dignity *oneself*,

independently of the moral law, through which self-love is made "lawgiving and the unconditional practical principle" [*CPrR* 5: 73–74]. This is not just a naïve tendency to mistake seeming normative relevance from one's perspective with objective normative weight. It is rather the radical idea that something has objective normative significance *because it is what one wills subjectively* – that one has a unique authority or standing to create reasons for acting independently of and unconstrained by the moral law. In the *Metaphysics of Morals*, Kant will add that it includes an important *interpersonal*, or as I will call it, "*second-personal*," dimension – not just that one can create reasons that others should take into account (on pain of irrationality, say), but that one has a unique authority to make claims and demands of others that they may not make of one, that they are answerable or accountable to one, but not *vice versa:* "lack of modesty in one's claims to be respected by others is *self-conceit* (*arrogantia*)" [*MM* 6: 462]. But we shall come to that.

A person with "self-love" confuses what she desires with the objectively desirable. In desiring that her thirst be relieved, for example, she takes it that there are reasons, in principle for anyone, to desire this state, that is, *her* thirst being relieved, and bring it about. But she doesn't think that others should do this because, that is, for the reason that, this is what she wants or wills. Rather, in wanting it, it just seems to her that such reasons do or must exist.[14] Self-conceit, on the other hand, is the idea that one's own will *is* a source of normative reasons (and is so uniquely). A thirsty person with self-conceit will take it that others have reasons to relieve his thirst *because this is what he wills or wants* (though he would have no such reasons if roles were reversed).

Self-conceit is the fantasy that one has a normative standing that others do not have to *dictate* reasons to oneself and others just because of who or what one is. The idea is not that one has the kind of authority that an especially good advisor does, that one sees better than others reasons that are there anyway. (Although it might involve this thought, if one took that to justify a special status also to create reasons by one's will.) It is rather the fantasy that one has a fundamental "lawgiving" standing

14 This is sometimes called the "backgrounding" of desire. On this important feature of practical reason, see, for example, Stephen Darwall, 1983, *Impartial Reason*, Ithaca, NY: Cornell University Press; Philip Pettit and Michael Smith, 1990, "Backgrounding Desire", in: *The Philosophical Review*, vol. 99, pp. 565–592; T. M. Scanlon, 1998, *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, pp. 41–55; and Jonathan Dancy, 2000, *Practical Reality*, Oxford: Oxford University Press.

that others simply don't have – as if one were King or God.[15] This is far from an innocent illusion, although Kant follows Rousseau in thinking that it is depressingly expectable whenever social comparison engenders *amour propre.*[16]

It follows that the moral law cannot simply curtail self-conceit or keep it in its place; it must "strike it down." It must declare "null and quite unwarranted" any "claims to esteem for oneself that precede accord with the moral law" [*CPrR* 5: 73]. We shouldn't be thrown off by Kant's use of 'esteem' here. The moral law must supplant self-conceit's presumptuous *authority*, and this is an object of recognition rather than appraisal respect (which responds to merit). To feel the authority of the moral law just is to sense that any conflicting claims to authority, such as that involved in self-conceit, are no better than puffery and pretense. And this experience is humiliating to the self-conceited agent.[17]

Now that a morally good agent who respects the moral law in the practical sense and takes it as sufficient reason to act, inclination costs to the contrary notwithstanding, will feel its humbling authority seems unproblematic. But Kant also has an interesting account of how we can come to feel respect for morality even when we are susceptible to self-conceit. To Fontenelle's remark that "I bow before an eminent man, but my spirit does not bow," Kant adds that "before a common humble man, in whom I perceive uprightness of character" greater than my own, my spirit does indeed bow. The encounter with unpretentious virtue can disarm the exceptionalist pretensions of self-conceit. "*Respect* is a *tribute* that we cannot refuse to pay to merit, whether we want to or not; we may indeed withhold it outwardly but we still cannot help feeling it inwardly" [*CPrR* 5: 77]. There is a fascinating point in the background here that can come out most clearly only when we fully appreciate the inter-

15 "Arrogance (*superbia* and, as this word expresses it, the inclination to be always *on top*) is a kind of *ambition* (*ambitio*) in which we demand that others think little of themselves in comparison with us" [*MM* 6: 465]. For a fascinating discussion of the role of the "wish to be God" in Kant's philosophy generally, see Susan Neiman, 2002, *Evil in Modern Thought: an Alternative History of Philosophy*, Princeton, NJ: Princeton University Press, pp. 57–84.

16 For an insightful account of Kant's ethics that stresses the role of self-conceit, see Wood, 1999.

17 Arguably, all of us indulge in this fantasy in some way or other. There are fascinating issues here concerning the interaction with Kant's views on radical evil in *Religion Within the Limits of Mere Reason* that I cannot get into here. For a fascinating discussion, see Robert Gressis, *Kant's Theory of Evil: An Interpretation and Defense*, 2007, University of Michigan, Ph.D. Diss.

personal dimension of respect and self-conceit, which Kant highlights in *The Metaphysics of Morals.* It is our "unsocial sociability" and interactions with one another that provide the context within which *amour propre* and self-conceit naturally arise. The exceptionalist fantasy only makes clear sense against the background of some felt pressure to justify ourselves *to* one another, that others have some *claim* to our doing so. Self-conceit is a kind of defense. But the need to defend oneself can be disarmed when the focus is off ourselves and we are simply attending to others, especially someone, "a common, humble man" who lacks any pretensions that might directly challenge our own. We see his virtue with undefended eyes and cannot help feeling the esteem it naturally engenders. The felt "tribute" we cannot refuse to pay to merit is, so far, a form of appraisal respect. But it also engenders recognition respect for the moral law, since "[h]is example holds before me a law that strikes down my self-conceit" [*CPrR* 5: 77]. He serves, among other things, as an "existence proof" of what Kant calls the moral law's "practicability" that can deflate exceptionalist fantasies born of self-serving cynicism, as when we let ourselves off the hook by saying or thinking to ourselves that compliance would involve unreasonable sacrifice, or that others don't comply.

Kant's focus in the second *Critique* is squarely on pure reason's capacity to determine the will, so we find there no extended discussion of dignity or the doctrine of rational nature as an end in itself to compare with the treatment of these topics in the *Groundwork.* What Kant says, however, along with the role he there gives respect in empirical psychology, seems consistent both with familiar interpretations and with the view that dignity is identical with merit or depends upon it, which we saw expressed in the passages we considered from the *Groundwork.* Kant's brief remarks about the "inviolability" of *humanity*, specifically, that a rational person "is not to be subjected to any purpose that is not possible in accordance with a law that could arise from the will of the affected subject himself" [*CPrR* 5: 87], could refer to a purely formal constraint posed by the formal end that Kant says rational nature "sets itself" of respecting the moral law [*G* 4: 437].

The first place we clearly find the thesis that all rational persons have a dignity regardless of their merit that warrants respect for them of a kind that differs from appraisal (merit) respect is in *The Metaphysics of Morals.* And here we find also an idea that will be at the absolute heart of the ideal of equal dignity that Kant bequeaths to modern moral thought, namely, that every rational person has a fundamental authority to make claims and demands of one another as an equal and, specifically, that all

have a claim to respect and, consequently, a corresponding duty of respect. In effect, Kant there holds the view that Rawls famously expressed by saying that persons are "self-originating source[s] of valid claims."[18]

III. Respect and the Duty of Respect in the Metaphysics of Morals

III.1. Equal dignity

Consider the following remarkable passage from the Section III ("On Servility"), Chapter III of Book I of the *Tugendlehre.* Kant is here discussing perfect duties to oneself "merely as a moral being."

> But a human being regarded as a *person*, that is, as the subject of a morally practical reason, is exalted above any price … he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. He can measure himself with every other being of this kind and value himself on a footing of equality with them.
>
> Humanity in his own person is the object of the respect which he can demand from every other human being, but which he must also not forfeit [*MM* 6: 434–435].

There are several things to emphasize about this passage. First, and most obviously, respect for an *authority* to claim or demand something is no form of respect for merit; it is not appraisal respect. The only sense in which merit can "demand" respect is by spontaneously giving rise to the esteem that constitutes appraisal respect. What authority calls for, by contrast, is *recognition*, and this we give and express by regulating our conduct, by how we conduct ourselves in relation to it. Second, the authority *to claim or demand* differs also from authority of other kinds in that it is, in a sense I shall explain presently, irreducibly *second-personal.* The intuitive idea can be put simply. Claims and demands are made or addressed to an addressee. So, third, the respect in which recognition of such an authority consists is also irreducibly second-personal; it involves *acknowledgment.* Consider the difference between this *second-personal authority*, as I shall call it, and epistemic authority. Respect for someone's judgment or knowledge is also a form of recognition, but epistemic

18 John Rawls, 1980, Kantian Constructivism in Moral Theory, *The Journal of Philosophy,* vol. 77, p. 546.

authority does not, as such, call for acknowledgment. If I overhear sage advice while eavesdropping, I can respect the advisor's wisdom in my private practical reasoning, simply by acting on the advice. Of course, if someone makes a *claim* to knowledge, then we do have something second-personal. The only way of respecting such a claim is with a second-personal response, either by directly acknowledging it, or by evidently relying on his claims in a way that is common knowledge. If I simply rely on his knowledge privately in my own reasoning, then, although I respect his epistemic authority, I do not thereby respect his claim to it, that is, his authority *to* claim such recognition of me. The former simply concerns his knowledge; the latter, his standing to enter into essentially interpersonal relations and reasoning in various ways.

Kant here treats an authority to make claims, specifically, to claim respect, as part of the dignity of persons. Dignity is "that by which [a person] exacts *respect* for himself from all other rational beings in the world" [*MM* 6: 435]. "Exacting" is a form of "demanding" as the first sentence of the next paragraph (quoted above) makes clear. And this is the fourth point: dignity includes a second-personal authority, specifically, the authority to make a claim for respect on "all other rational beings." Now, as I have said, any form of authority calls for respect, but second-personal authority of any kind calls for a distinctively second-personal form of respect; since it invariably makes a putatively valid claim of some sort of others, it invariably calls for respect of the authority to make that claim. But not every kind of second-personal authority *demands* respect. When, to return to our example, someone makes a claim to knowledge on some matter, although in so doing he makes a claim to respect of the authority to make that claim, he does not necessarily demand such respect – he may have no such authority. Clearly, however, Kant is taking it to be central to the dignity of persons that they have the authority to demand respect. Finally, fifth, Kant also plainly says that the authority to demand respect puts us on "a footing of equality" with one another. The respect we have the authority to demand is thus respect *as an equal.*

When we considered Kant's account of self-conceit in the second *Critique* we noted that Kant took it to involve the pretense of a law-giving standing in one's own will that one does not recognize in others. In that work, however, there was nothing essentially second-personal about self-conceit. It is one thing to think that others have a reason to do something because this is what one wills, and another to think that one has a claim to others' doing what one wills or that one has the authority to demand

that they do so. In the *Tugendlehre*, however, Kant characterizes self-conceit in distinctively second-personal terms. "Self-conceit," he says, is "lack of modesty in one's claims to be *respected* by others." And its antidote is recognition of (equal) authority to demand respect as an equal: "Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other" [*MM* 6: 462].

Presently, we shall consider the duty of respect. But first I want to say more about the idea of second-personal authority that I am claiming is central to the dignity of persons as Kant understands it in *The Metaphysics of Morals.* Suppose someone is standing on your foot. Consider three different kinds of reasons there might be for him to remove his foot from on top of yours.[19] One might simply be the badness of your being in pain and the fact that he could bring about a better state of the world if he were to move his foot. But suppose, second, that you and he hold a norm of conduct according to which one should not step on others' feet and should remove one's foot promptly should one find it on top of someone else's. This would be a different reason than the first. The first would be *agent-neutral* reason for removing his foot.[20] The reason would not be essentially *for him* as the agent causing another person pain. It would exist, most fundamentally, for anyone who is in a position to effect your relief and *therefore* for him, since he is well placed to do so. Since, however, the second reason derives from an *agent-relative* norm (that is, a norm that is reflexive and where the reason-description makes ineliminable back reference to the agent), it is an agent-relative reason, unlike the

19 I discuss these matters at much greater length in Stephen Darwall, 2006, *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press, on which I here draw.

20 Agent-neutral reasons contrast with agent-relative reasons, whose formulation includes an ineliminable reference to the agent for whom they are reasons (like "that it will keep a promise I made," "that it will avoid harm to others (i.e., people other than me)", and so on). Agent-neutral reasons can be stated without such a reference: "that it would prevent some pain from occurring to someone (or some being)." On the distinction between agent-relative (also called "subjective" or "agent-centered") and agent-neutral (also called "objective") reasons, principles, values, etc., see Thomas Nagel, 1970, *The Possibility of Altruism*, Oxford: Clarendon Press; Samuel Scheffler, 1982, *The Rejection of Consequentialism*, Oxford: Clarendon Press; Derek Parfit, 1984, *Reasons and Persons*, Oxford: Oxford University Press; Stephen Darwall, 1986, "Agent-Centered Restrictions from the Inside Out", in: *Philosophical Studies*, vol. 50, pp. 291–319.

first. Despite the fact that violations of the norm are all equally criticizable, or bad, the norm counsels that one not do what would violate it *oneself.* If, for example, the person on your foot could prevent two people from comparable foot-mashings by the shocking spectacle of keeping his foot firmly planted on yours, the norm would not counsel him to do so.

Consider now the idea, not just that the agent has a reason (of whatever weight or priority) not to step on others' feet, but that persons have the *authority to demand* that they not step on one another's feet. This adds a new reason of a fundamentally different kind. We might think, for example, that all individuals have the authority to demand that others not step on their feet without their consent, because we have the right not to have our feet stepped on. Alternatively, or in addition, we might think that stepping on people's feet is something we have the authority to demand, as equal members of the moral community (the realm of ends), that persons not do to one another, and, therefore, that we are answerable or accountable for not doing so to one another, and that it is wrong, contrary to a moral obligation, to do so.

This third kind of reason not to step on others' feet is agent-relative like the second also, but it is an additional reason; it is the (further) reason that others (or you) validly claim or demand that we not step on others' feet without their consent. I might accept an agent-relative norm of whatever stringency, generating reasons for acting of whatever weight or priority, that proscribes stepping on others' feet without accepting that this is anything they have the authority to claim or demand.

What is distinctive about reasons of this third kind is their essential connection to second-personal address. Someone can sensibly accept this third reason for moving his foot, one embodied in a claim or demand, only if he also accepts the authority *to* demand this of him (second-personally). That is just what it is to accept something *as a valid claim or demand.*[21] And if he accepts that you can demand that he move his foot, he must also accept that you will have grounds for complaint or some other form of accountability-seeking response if he doesn't. Unlike the first and second reasons, this third reason is second-personal in the sense that, although the first and second are conceptually independent of the second-personal address involved in making claims and holding persons responsible, the third is not.

21 There are, of course, ways of accepting demands, say out of self-interest in a negotiation, that are different from accepting something as a valid demand.

Second-personal reasons are invariably tied to a distinctively second-personal kind of *practical authority:* the authority to make a demand or claim. Making a valid claim or a demand always presupposes the authority to make it and that the duly authorized claim creates a distinctive reason for compliance (a second-personal reason). Moreover, these notions all also involve the idea of responsibility or accountability. The authority to demand implies, not just a reason for the addressee to comply, but also his being responsible to the addresser (at least) for doing so. Conversely, accountability implies the authority to hold accountable, which implies the authority to claim or demand, which is the standing to address second-personal reasons. These notions – second-personal authority, valid claim or demand, second-personal reason, and responsibility to – therefore comprise an interdefinable circle; each implies all the rest. Moreover, I contend, there is no way to break into this circle from outside it. Propositions formulated only with normative and evaluative concepts that are not already implicitly second-personal cannot adequately ground propositions formulated with concepts within the circle.

There is an important difference between the idea of an authoritative (second-personal) claim or demand, on the one hand, and that of an authoritative or binding norm in the familiar sense of a valid ought that entails genuine normative reasons, on the other. So when Kant says that the dignity of persons includes the authority to exact or demand respect he implies that it includes an authority to make claims and demands of one another, and so hold one another responsible, as equals. The idea that to be a person is to be a "self-originating source of valid claims" goes beyond any normative thesis about the existence of reasons of whatever normative priority or weight for persons to treat one another in various ways. It entails the further proposition that persons are mutually accountable for this treatment, that they must be able to justify their treatment of each other *to* one another.

Kant's *Tugendlehre* claim that dignity includes an equal second-personal authority thus goes beyond anything we can find in either the *Groundwork* or *The Critique of Practical Reason* even on the traditional interpretation. Even the idea that persons have an inviolable dignity in the sense that they may not be treated as means does not entail that this is something that persons can themselves demand or for which we are answerable to one another.[22]

22 Unless, of course, we interpret 'may not' to entail 'responsible to one another not

III.2. The duty of respect

I turn now to the duty of respect, both to oneself and to others. Kant discusses both under the heading of ethical duties, the first to himself "merely as a moral being" (I.i.2), and the second, to others "merely as human beings" (II.i.3). But how are we to understand the respect for oneself or for others that it is one's duty to have? As I mentioned earlier, we can be misled by Kant's saying that we must respect persons even if they "mak[e themselves] unworthy of it" in their deeds [*MM* 6: 463]. Since the respect that deeds can justify is appraisal respect (for their merit), this might encourage the thought that we should somehow show or have such (appraisal) respect for someone in spite of what they actually do. That this cannot be what Kant means, however, is entailed both by Kant's overall project in the *Tugendlehre* and by what he says there about the respect.

Duties of virtue, "ethical duties," as Kant calls them, are concerned not with the "legality" of actions (that is, with which acts are required considered independently of motive) but with which ends or maxims we should have – with "internal" rather than "external lawgiving" [*MM* 6: 394–395]. An ethical duty of respect, whether for oneself or for others, consequently, is a duty to have a certain end or maxim. If the duty of respect were a duty to show or have appraisal respect (as for someone's merit, whether deserved or not), then it would have either to be a duty to act as if one had it, regardless of whether one did, or somehow to have it internally whether or not it is a fitting response to the person's believed or actual merit. But a duty simply to act in a certain way (regardless of motive) is not an ethical duty, and having appraisal respect for someone is not anything, like a maxim, that one can adopt at will. This is simply the flip side of Kant's dictum that (appraisal) "respect is a tribute that one cannot refuse to pay to merit" since we "cannot help feeling it inwardly" for someone we see as virtuous. If we regard someone as vicious, then we cannot help feeling appraisal contempt for him, or at least, find ourselves spontaneously unable to feel respect.

Moreover, Kant says explicitly that the respect called for by the duty of respect is not *reverentia* (a feeling of respect, whether the phenomenal aspect of respecting the moral law or felt respect for merit); it is "respect

to'. But then we would have to ask what within Kant's framework in the *Groundwork* or *The Critique of Practical Reason* will license or legitimate the claim thus interpreted.

in the practical sense," or *observantia.* It is, he says, "the *maxim* of limiting our self-esteem by the dignity of humanity in another person" [*MM* 6: 449]. Again, as we noted with a similar passage from *The Critique of Practical Reason* above, we should not be thrown off by "self-esteem." Esteem in the usual sense of a response to merit is not the sort of thing that one *can* limit by adopting a maxim. What Kant must mean, I believe, is that we must limit our *claims* to authority by the equal (second-personal) authority that any person has simply as such. This we owe to others. But we owe no less to ourselves. Respect for oneself, for one's own dignity as a person, requires that one respect that same equal authority in oneself.

The latter is evident in the perfect duty to oneself "as a moral being" not to be "servile" [*MM* 6: 434–435]. Respect for one's dignity is not only something one "can demand from every other human being," but something one "must also not forfeit" [*MM* 6: 435]. So one must "be no one's lackey" and not "let others tread with impunity on [one's] rights" [*MM* 6: 436]. To acquiesce in others' arrogance, in their immodesty in "claims to be respected by others," is to fail in proper respect for oneself. This follows from the interactive or second-personal nature of claiming. A claim to authority over another person is addressed to an addressee over whom authority is claimed. Failure to challenge the claim, to stand up for oneself, can be taken as acquiescence, as implicitly acknowledging it. Similarly, a right, in its very nature, is something we can validly claim. As Joel Feinberg pointed out, "it is claiming that gives rights their special moral significance . . . [h]aving rights enables us to 'stand up like men,' to look others in the eye, and to feel in some fundamental way the equal of anyone."[23] The very idea of a right involves that of an authority or standing to claim or demand something, a place to stand in asserting a claim or addressing a valid demand. Kant invokes the same metaphor himself when he says that our dignity enables a person to "value himself on a footing of equality with them" [*MM* 6: 435].

Respect for oneself as a person thus involves insisting on one's equal second-personal authority to make claims and demands *as* a person, where part of what one can demand of oneself and others is *respect for this (equal second-personal) authority.* And since this equal authority is fundamentally second-personal, so also is the respect that it demands; it is *acknowledgment* rather than any thing that can be realized simply in

23 Joel Feinberg, 1980, "The Nature and Value of Rights," in: *Rights, Justice, and the Bounds of Liberty*, Princeton, NJ: Princeton University Press.

an agent's private practical reasoning. Just as the duty of respect to others requires, as Kant says, our being "under an obligation to acknowledge, in a practical way, the dignity of humanity in every other human being," so also must we challenge others' attempts to deny it to us [*MM* 6: 579].

This brings us to the duty of respect to others, and to a contrast that Kant draws there between the duty of respect and the duty of love. Together these exhaust ethical duties to others "merely as human beings" [*MM* 6: 448]. There are two points we should bear in mind in interpreting Kant's idea. The first is that Kant is here talking about *ethical* rather than juridical or legal duties; hence the ethical duties of love and respect both concern duties to have a specific maxim. The maxim called for by the duty of love is "to make others' ends my own (provided only that these are not immoral)" [*MM* 6: 450]. And the maxim of respect, again, is to acknowledge others' dignity "in a practical way" [*MM* 6: 462]. That they are both duties to have a certain maxim or end is common to the duties of love and respect as ethical duties. What distinguishes them, second, is not just their respective maxims, but also that although respect is something we owe to all others, love is not. Performance of the duty of love is thus "*meritorious* (in relation to others)" [*MM* 6: 448]. There is actually a delicate point here signaled in Kant's parenthetical addition of "in relation to others." The idea is not that we warrant special moral esteem when we act out of the duty of love, since there is, after all, an imperfect duty of love to make the ends of others our own. We may be doing no more than is already our (imperfect) moral duty. The point is that we are doing something that goes beyond any duty we owe *to the specific individuals we benefit*, since they have no ground to *expect* our benevolence. No one can claim love. So since acts of love go beyond anything individuals have any claim to, they then "put others under obligation" [*MM* 6: 448]. They give rise to a duty of gratitude to reciprocate [*MM* 6: 454–455].

The duty of respect is, however, owed to all others. It is something that, as we have seen, all persons have the authority to claim or demand. Consequently, observance of the duty of respect to others "does not result in obligation on the part of others" [*MM* 6: 448]. Kant treats respect and love as two "great moral forces" that govern goings on in the intelligible realm much as gravitation does in the physical world – "the principle of mutual love admonishes them [persons] to *come closer* to one another; that of the *respect* they owe one another to keep themselves *at a distance*

from one another" [*MM* 6: 449].[24] We shouldn't read too much into this contrast, however. It does not entail, for example, that we only owe others "negative" rather than "positive" duties. *Whatever* we owe others mobilizes the duty of respect. So, for example, when others have benefited us by their love, Kant believes that we owe them the duty of gratitude, and when we discharge this duty and reciprocate, we do so out of respect for their authority, since we do no more than the benefactor has claim to expect [*MM* 6: 454]. Similarly, if we fail to show gratitude, this is a failure of respect, not of love, Kant thinks. We fail to recognize the person's valid claim to reciprocation and so his standing to claim it.

Kant distinguishes three specific kinds of disrespect that the duty of respect requires us to avoid: arrogance, defamation, and ridicule. Arrogance, again, is self-conceit, immodesty in "one's claims to be respected by others" [*MM* 6: 462]. Whenever we knowingly fail to discharge duties we owe to others we effectively privilege ourselves over them and give ourselves a privilege of claim that we deny them by our action. Kant insightfully remarks that "arrogance is, as it were, a solicitation on the part of one seeking honor for followers, whom he thinks he is entitled to treat with contempt" [*MM* 6: 465]. "Who does he think he is?," we say of the person who prevails on others' patience and forbearance. "He acts as if he expects others to accommodate themselves to his wishes." And in so doing, this person in effect makes a claim on others, presumes thereby the standing to do so, and effectively "demand[s] that others think little of themselves in comparison with [him]" [*MM* 6: 465]. Since duties owed *to* others are what they can legitimately claim deriving from their authority as "self-originating sources of valid claims," in Rawls's phrase, failures to discharge them are failures of respect. So they also violate the duty of respect.

But there are ways of failing to respect others' dignity as persons without treating their (other) valid claims with contempt. Kant distinguishes two: defamation and ridicule. We *defame* others in the sense Kant has in mind when we engage in fault-finding, rumor-mongering, and the like so that others (of oneself) might be encouraged not to treat our targets as full moral persons or in ways that tend "to cast a shadow of worthlessness over our race itself," feeding cynicism and misanthropy that undermine moral community [*MM* 6: 466].

---

24 See Marcia Baron, 2002, "Love and Respect in *The Doctrine of Virtue*", in: Mark Timmons (ed.), *Kant's Metaphysics of Morals*, Oxford: Oxford University Press; for a critical discussion of this aspect of Kant's views.

Similarly, we disrespect others by *ridicule* when we mock or "expose others to laughter" and "make their faults the immediate objects of [our] amusement" [*MM* 6: 467]. Kant makes an interesting distinction between ridicule and "banter" that "makes fun" of others' "peculiarities that only seem to be faults but are really marks of their pluck in sometimes departing from the rule of fashion" [*MM* 6: 467]. Nonetheless, "holding up to ridicule a person's real faults, or supposed faults as if they were real," including their moral faults, "in order to deprive him of the respect he deserves," is a violation of the duty of respect [*MM* 6: 467]. This suggests a deep point about the role of respectful address in holding people responsible. As P. F. Strawson noted in "Freedom and Resentment," when we hold others responsible, we take up an engaged, or as I have been putting it, second-personal, attitude toward them that addresses a claim or demand but that also expresses respect. As Strawson put it, we "continue to view [the other] as a member of the moral community; only as one who has offended against its demands."[25] We find the same point in Kant's idea that punishment is a way of respecting the dignity of humanity in the wrongdoer [*MM* 6: 362–363]. To hold someone responsible for wrongdoing as a person is to demand respect *from* them, but in a way that simultaneously expresses respect *for* them, as an equal member of the moral community.[26] When, however, we ridicule people for moral faults, we make them *objects* of ridicule, that is, the object of a third-personal regard that does not engage and treat them as equal persons. Ridiculing them for their moral faults deprives the moral community (and them) of the accountability that all can claim and tends to diminish the mutual responsibility on which the moral community (the realization of the kingdom of ends) itself depends.

IV. Conclusion

I have been arguing that, as Kant holds in *The Metaphysics of Morals* that the dignity of persons is a second-personal authority "by which" we can "exact" or demand respect from one another as equal moral persons, so also must the respect we can demand be a second-personal acknowledgment of this authority. Elsewhere I have argued that this equal

25 P. F. Strawson, 1968, "Freedom and Resentment", in: *Studies in the Philosophy of Thought and Action*, London: Oxford University Press, p. 93.

26 See *The Second-Person Standpoint* for the development and defense of this idea.

fundamental second-personal authority can be read back into Kant's earlier ethical writings in ways that can help fill significant lacunae in those works.[27] If to be a person is to have equal standing in interpersonal relations, to have an equal authority to make claims of one another, then the moral law is not just what we are committed to within our own individual practical reasoning, it also makes possible "a systematic union of various rational beings through *common* laws," as Kant puts it in the *Groundwork* ([*G* 4: 433], emphasis added). Bringing in equal second-personal authority makes *mutual accountability* to one another as free and rational persons an essential aspect of moral obligation.[28]

The main philosophical problem that Kantian ethics faces is to show that its central moral ideas of the dignity of persons, the Categorical Imperative, the unqualified goodness of the good will, autonomy of the will, and so on, are not illusory fantasies – "phantom[s]" or "chimerical ideas," as Kant puts it at the end of Chapter II of the *Groundwork* [*G* 4: 444–445]. In the *Groundwork*, Kant famously argues that any rational agent must presuppose autonomy in order to deliberate intelligibly, and that this assumption brings the other central moral ideas in its train. By the time of the second *Critique*, however, Kant seems to have given up this strategy, arguing that it is only through our awareness of being bound by the moral law that we can see that we are autonomous [*CPrR* 5: 4n]. In my view, Kant was right to abandon the *Groundwork*'s strategy of arguing that autonomy of the will and, therefore, the other moral ideas that follow from it, are essential presuppositions of the practical point of view. I agree also with the critics that more recent attempts to derive these central moral ideas from the assumptions of the practical standpoint do not work either.[29] In my view, no attempt to derive autonomy of the will and the moral law from the first-person practical point of view alone can possibly work. Moreover, it is the lesson of what Kant says about the dignity and respect in *The Metaphysics of Morals* that, even if such an

27 In *The Second-Person Standpoint*, especially Chapters 9 and 12.

28 For an excellent discussion of Kant's ethics from this perspective, see Christine Korsgaard's "Creating the Kingdom of Ends", in: *Creating the Kingdom of Ends.*

29 See, e.g., Christine Korsgaard, 1996, "Two Distinctions in Goodness, and Kant's Formula of Humanity", in: *Creating the Kingdom of Ends*; Christine Korsgaard, 1996, *The Sources of Normativity*, Cambridge: Cambridge University Press; and Allen Wood, *Kant's Ethical Thought.* For criticism, see, e.g., Rachel Cohon, 2000, "The Roots of Reason", *The Philosophical Review*, vol. 109, pp. 63–85; and Samuel Kerstein, 2001, Korsgaard's Kantian Arguments for the Value of Humanity", in: *Canadian Journal of Philosophy*, 31, pp. 23–52.

argument were to succeed in its own terms, showing that we always have conclusive reason to comply with the moral law, it still would not establish that compliance with the moral law and respect for the dignity of persons is something that person have themselves the *authority to demand* (second-personally). There simply is no way to establish these without invoking concepts that are already within the circle of irreducibly second-personal notions that I mentioned before. It follows, I believe, that the only way properly to locate and establish these key Kantian moral ideas is from within the second-person standpoint. I have argued elsewhere that this can indeed be done.[30] The presuppositions of intelligible second-personal practical thought commit an agent to respect for the dignity of persons and autonomy of the will, and, moreover, these claims can be vindicated within a comprehensive theory of practical reason that includes the second-person point of view.[31] What I have argued here is that the fundamental idea underlying this line of thought, the equal second-personal authority of all persons, is one that Kant himself articulates in *The Metaphysics of Morals.*

30 In Chapter 10 of *The Second-Person Standpoint.*
31 In Chapter 11 of *The Second-Person Standpoint.*

Treating Oneself Merely as a Means

Samuel J. Kerstein

Each of us is morally required not to lie, not to kill himself, and not to defile himself by lust (e.g., masturbate), according to Kant.[1] Kant attempts to derive these "perfect duties to oneself" from the Formula of Humanity: "So act that you treat humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" ([*G* 4: 429], italics omitted).[2] He tries to show that if

1 The other main "perfect duties to oneself" that Kant highlights are those not to stupefy oneself by the excessive use of food and drink, not to engage in greedy avarice, and not to be servile. Following his discussion of these six duties [*MM* 6: 422–437], Kant entitles a section "On the Human Being's Duty to Himself as his own Innate Judge" [*MM* 6: 437–440]. He does not there derive a particular duty to oneself, but rather offers a detailed description of the conscience. At the end of the division "Perfect Duties to Oneself" (Part I, Book I of the Doctrine of Virtue), Kant places an "Episodic Section" in which he discusses three further duties to ourselves, namely a duty not to develop a "propensity to wanton destruction of what is *beautiful* in inanimate nature" [*MM* 6: 442–443], a duty not to treat animals cruelly, and, finally, a duty to recognize all our duties as divine commands. It is odd that Kant discusses these duties under the division "Perfect Duties to Oneself." One might wonder, for example, why the duty not to treat animals cruelly is a duty to oneself, while the duty to sympathetic feeling is a duty to others. For Kant's rationale for the former duty is that a person's violent and cruel treatment of animals "dulls his shared feeling of their suffering and so weakens and gradually uproots a natural predisposition that is very serviceable to morality in one's relations with other people" [*MM* 6: 443].
All references to Kant are to the "Preussische Akademie der Wissenschaften" edition of his works. Under each entry, I specify the English edition I have consulted, each of which has academy edition page numbering in its margins. The translations I employ sometimes vary from those of these English editions: I use *Groundwork of the Metaphysics of Morals*, tr. Mary J. Gregor, in: Immanuel Kant: *Practical Philosophy*, 42–108. Cambridge: Cambridge University Press, 1996. *Critique of Practical Reason*, tr. Mary J. Gregor, in: Immanuel Kant: *Practical Philosophy*, 138–271. Cambridge: Cambridge University Press, 1996. *The Metaphysics of Morals*, tr. Mary J. Gregor, in: Immanuel Kant: *Practical Philosophy*, 363–603. Cambridge: Cambridge University Press, 1996.

2 Kant comes close to restating the Formula of Humanity at [*MM* 6: 395] and [*MM* 6: 462]. In derivations of four of the six duties he sets forth from [*MM* 6:

an agent acts contrary to the duties, then he treats himself merely as a means and thereby violates this principle. In order to understand fully Kant's rationale for the duties, we need, therefore, to pinpoint what it means to treat oneself merely as a means.

When Kant writes of treating oneself or others merely as means, he is referring to treating beings with "humanity" in this way. "Humanity" does not refer to the class of human beings, but rather to a set of capacities. Kant tells us that "the capacity to set oneself an end – any end whatsoever – is what characterizes humanity (as distinguished from animality)" [*MM* 6: 392]. So at the very least, if a being has humanity, then it has the capacity to set ends. Kant, it seems, uses "humanity" interchangeably with "rational nature" (see, e.g., [*G* 4: 439]). In doing so he suggests that having humanity involves having a whole set of rational capacities. Among them is the capacity to act on maxims and hypothetical imperatives, as well the capacity to act autonomously, that is, (roughly) to conform to self-given moral imperatives purely out of respect for these imperatives.[3] In what follows, references to treating oneself or others in some way are shorthand for treating the humanity in oneself or others in this way.

Philosophers have paid relatively little attention to the task of discerning what, according to Kant, it means to treat oneself merely as a means, focusing more on specifying what it means to treat others in this way. Kant gives his most thorough indication of what it means to do the latter in the *Groundwork of the Metaphysics of Morals* when he tries to derive from the Formula of Humanity a duty not to make false promises:

422] to [*MM* 6: 437], Kant obviously relies on this principle. In his derivation of two others, namely the duty not to stupefy oneself by the excessive use of food or drink and the duty not to engage in greedy avarice, the role of the Formula of Humanity is less clear. But his derivation of the duty not to stupefy oneself at least appears to stem from the Formula of Humanity. For Kant's objection to doing so seems to be that it severely diminishes one's capacity to exercise one's rational capacities. Perhaps Kant holds that diminishing this capacity is, other things being equal, incompatible with holding humanity to be an end in itself. If Kant appeals at all to the Formula of Humanity in his derivation of the duty not to engage in greedy avarice, it is in the rather obscure fifth paragraph of section 10 [*MM* 6: 432–433].

3 Here I am following Thomas E. Hill Jr., 1992, "Humanity as an End in Itself," in: *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca: Cornell University Press, pp. 38–57, at pp. 28–41. For a slightly different account of what Kant means by humanity, see Allen W. Wood, 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press, pp. 118–120.

> He who has it in mind to make a false promise to others sees at once that he wants to make use of another human being *merely as a means*, without the other at the same time containing in himself the end. For, he whom I want to use for my purposes by such a promise cannot possibly agree to my way of behaving toward him, and so himself contain the end of this action [*G* 4: 429–430].

According to one influential interpretation, namely that of Onora O'Neill, Kant here implies that an agent treats another merely as a means and thus wrongly if in his treatment of the other the agent does something to which the other cannot consent.[4] O'Neill implies that an agent can consent to (or, equivalently, agree to) a course of action only if it is possible for him to dissent from it. It is possible in the relevant sense for someone to dissent from a course of action, she says, only if he "can avert or modify the action by withholding consent and collaboration".[5] According to O'Neill, if an agent deceives or coerces another, then the other's dissent is "in principle ruled out," and thus so is his consent.[6] Suppose, for example, that an auto mechanic makes a false promise to a customer to have his repair done by 5:00 PM. The customer does not really have the opportunity to dissent to the mechanic's action. For he does not know what her action is, namely one of lying to him about when his car will be ready. Deceit, along with coercion, are central cases of some agents treating others merely as means.

However plausible this reading of treating others merely as a means might be, it does not help us to understand treating oneself in this way.[7]

---

4 Onora O'Neill, 1989, "Between Consenting Adults," in: *Constructions of Reason*, Cambridge: Cambridge University Press, pp. 105–125, at p. 113. Christine Korsgaard seems to agree with Onora O'Neill on this point. "The question whether another can assent to your way of acting," writes Korsgaard, "can serve as a criterion for judging whether you are treating her as a mere means", Christine Korsgaard, 1996, "The Right to Lie: Kant on Dealing with Evil", in: *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, pp. 133–158, at p. 139. More precisely, O'Neill suggests that an agent treats another merely as a means and thus wrongly if in his treatment of the other he acts *on a maxim* to which the other cannot consent. It is notoriously difficult to specify what Kant means by a maxim, and for the sake of simplicity I do not invoke maxims here. My not doing so does not so far as I can tell affect the substance of what follows.

5 O'Neill, 1989, p. 110.

6 O'Neill, 1989, p. 111.

7 In my opinion, Kant might be committed to an account of treating others merely as means along the lines of the one O'Neill attributes to him. But, as I argue in "Treating Others Merely as Means" (manuscript), I do not believe this account to be philosophically plausible.

Kant holds that an agent treats himself merely as a means if he kills himself, masturbates, or lies. But in these cases, the agent is able to consent to the way he treats himself; he can avert or modify this treatment. All of these actions are, of course, in violation of the categorical imperative, according to Kant. But to "satisfy the categorical command of morality is within everyone's power at all times," he tells us [*CPrR* 5: 36–37]. In short, since agents are free, they would in Kant's view never be unable to consent to their treating themselves in a morally impermissible way.[8] So in order to understand Kant's conception of an agent's treating himself merely as a means it does not suffice to consider O'Neill's account of what, according to Kant, it means to treat others in this way.

This paper attempts to shed light on what, according to Kant, it means to treat oneself merely as a means. Since Kant relies on this notion in connection with his discussion of duties to oneself, I discuss briefly his conception of these duties (Section I). In Section II I consider a simple answer to the question of what it means to treat oneself merely as a means, namely that doing so just amounts to failing to treat oneself as an end in himself, that is, as something that is unconditionally and incomparably valuable. This answer does not allow us to understand how Kant actually employs the notion of treating oneself merely as a means, I argue. Section III is devoted to developing a different interpretation of this notion – an interpretation that stems from ideas implicit in the false promising passage cited above. In Section IV, I use this interpretation in an attempt to illuminate Kant's efforts to derive duties not to murder oneself, not to defile oneself by lust, and not to lie. But I do not offer anything approaching an interpretation of all of the arguments Kant

8 One might argue that there are rare cases in which an individual uses himself in some way yet cannot, in the relevant sense, consent to this usage. For example, an agent chooses not only to be brainwashed into believing that he'll die if he drinks a sip more of alcohol, but also into forgetting that he ever agreed to or underwent such a procedure. The agent treats himself in some way: he gets brainwashed in order to break his addiction to alcohol. Suppose that it is legitimate to think that the self who chose to get brainwashed on Monday is now, on Friday, treating his brainwashed self in some way: the former is doing something to the latter. On this supposition, the legitimacy of which Kant would, I think, deny, we can see how the agent's former self would be treating his current self merely as a means. For the current self cannot avert or modify the former self's action. He does not even know what this action is.

suggests for these duties. The final section highlights some difficulties with Kant's efforts to derive them.

I.

The duties that are such that in failing to abide by them one is treating oneself merely as a means are duties to oneself, according to Kant. It thus makes sense to consider briefly Kant's general discussion of duties to oneself. Unfortunately, this discussion is not particularly illuminating.

The notion of a duty to oneself contains an apparent contradiction, Kant tells us. It belongs to the concept of a duty that if an agent has one, then he is "passively constrained" [*MM* 6: 417] to conform to the duty. Yet it is self-contradictory to affirm that an agent is passively constrained if he is the very agent who imposes the duty. For if it is he who imposes the duty, then he can always release himself from it. But if he can always release himself from the duty, then, by definition, he is not passively constrained to conform to it. In short, it is self-contradictory to claim that it is the very same entity who imposes an obligation and who is under this obligation. Yet if we maintain there to be duties to oneself, then we are, it appears, committed to this claim.

Kant tries to show that in fact we are not committed to it. When we reflect, we realize that in our view the being who imposes obligation is not one and the same as the being who stands under it. Kant mentions two aspects of the human being: the sensible being (*homo phaenomenon*), that is, an animal with reason and the intelligible being (*homo noumenon*), that is, a free agent [*MM* 6: 418]. According to Kant, the *homo phaenomenon* cannot be put under obligation. But "a being endowed with inner freedom (*homo noumenon*), is regarded as a being that can be put under obligation" [*MM* 6: 418], he says. It would thus seem natural for Kant to affirm that it is not the very same being who imposes an obligation and who is under it. It is, rather, one aspect of an individual, the *homo phaenomenon*, who imposes an obligation and another aspect of this individual, the *homo noumenon* who is under it. So the notion of a duty to oneself is not self-contradictory.

But Kant does not affirm this. He says, rather, that the "human being thought in terms of his personality, that is, as a being endowed with inner freedom (*homo noumenon*), is regarded as a being that can be put under obligation and, indeed, under obligation to himself (to the humanity in his own person)" [*MM* 6: 418]. According to Kant, humanity is

constitutive of the *homo noumenon* (see, e.g., [*MM* 6: 239, 295] and Vigilantius "Notes on the lectures of Mr. Kant on the metaphysics of morals," [27: 627–628]). So Kant is maintaining here that the *homo noumenon* can be put under obligation to the *homo noumenon.* The *homo phaenomenon* drops out of the relation. The difficulty Kant initially raised remains: how, without contradiction, can one maintain that a human being – in particular a human being considered solely as *homo noumenon* – both imposes constraint (an obligation) on itself and is passively constrained?[9]

Kant develops a tool that might enable him to give a satisfactory answer to this question, although he does not explicitly use it in the context of discussing the apparent contradiction in a duty to oneself. This tool is his distinction between the will [*Wille*] and choice [*Willkür*]. Kant tells us that:

> Laws proceed from the will, maxims from choice. In man the latter is a free choice; the will, which is directed to nothing beyond the law itself, cannot be called either free or unfree, since it is not directed to actions but immediately to giving laws for the maxims of actions (and is, therefore, practical reason itself). Hence the will directs with absolute necessity and is itself subject to no necessitation. Only choice can therefore be called free ([*MM* 6: 226], italics omitted).

Perhaps, according to Kant's considered view, the *homo noumenon* itself has two aspects: free choice [*freie Willkür*], which is a capacity to act without being determined to do so by any sensible impulse (see, e.g., [*MM* 6: 213–214]) and the will [*Wille*], which is a capacity to set forth unconditionally binding moral laws. Free choice is constrained and the will does the constraining. If this interpretation is on target, Kant might be able to maintain that, despite appearances to the contrary, there is no contradiction in holding that the *homo noumenon* has duties to itself.

In any case, this is the interpretation I will assume. Strictly speaking, the being who has duties to itself is the *homo noumenon* – in particular its capacity of free choice; the being who promulgates these duties is a different aspect of the *homo noumenon*, namely the will. (The will promulgates them through presenting or "legislating" the moral law.) For simplicity's sake, in what follows I will not employ the term "*homo noumenon*," but will instead simply write of a person or of humanity.

---

9 Andrews Reath makes a similar point (2002, "Self-Legislation and Duties to Oneself," in: Timmons (ed.), *Kant's Metaphysics of Morals: Interpretive Essays*, Oxford: Oxford University Press, pp. 349–370, at p. 356).

II.

In the Doctrine of Virtue, Kant reaffirms his *Groundwork* conviction that humanity has dignity. It has "absolute inner worth," a value which is beyond any price and which exacts respect [*MM* 6: 434–435]. Violating perfect duties to oneself somehow involves a failure to give humanity, namely one's own, the respect it demands. In Section III, I sketch an interpretation/reconstruction of what it means for an agent to treat himself merely as a means. Based on the *Groundwork* passage on false promising quoted above, the interpretation lays out a detailed procedure for determining whether an agent is treating himself in this way. Examining this procedure will, I hope, help us to see precisely how, in violating his perfect duties not to kill himself, defile himself by lust, or lie, an agent fails to respect his rational nature.

But one might ask whether reconstructing such a procedure on Kant's behalf is really necessary. In deriving other perfect duties to oneself, namely those not to stupefy oneself with too much food or drink, not to be avaricious, and not to be servile, Kant does not rely on the notion of treating oneself merely as a means. So why get bogged down in developing a detailed procedure for determining whether an agent treats himself merely as a means? Why not, in the case of each of the perfect duties to oneself Kant enumerates, interpret him to be simply appealing to (what he takes to be) our view that acting contrary to them expresses disrespect for humanity?

A difficulty with this approach is that it can leave rather mysterious how Kant arrives at some of the duties. Consider, for example, a recent treatment of Kant's derivation of the duty not to lie. According to Lara Denis, Kant claims that "when we profess to be expressing our minds, intentionally speaking in a way that misrepresents what we think shows a lack of respect for our rational nature." It shows a lack of respect in that it "expresses an insufficient commitment to represent externally our rational nature".[10] For we represent our rational nature externally by speaking truthfully, not by lying.

This is a puzzling argument. Suppose that lying does express a lack of commitment to represent our rational nature externally.[11] What would constitute a sufficient commitment? Isn't refraining from representing

10 Lara Denis, 1997, "Kant's Ethics and Duties to Oneself," in: *Pacific Philosophical Quarterly*, vol. 78, pp. 321–348, at p. 331.

11 I'm not at all sure that it does. Who but a rational being is capable of lying?

our rational nature externally, at least on some occasions, compatible with respecting it as something of absolute inner worth? If lying always reveals an inadequate commitment to represent our rational nature externally, does remaining silent when we could be expressing our thoughts do so as well? Denis suggests that Kant's answer is no. But what is (or might be) his justification for this answer? There remains a gap between the idea that, in virtue of its special value, an agent must respect his humanity and the finding that lying involves his disrespecting it.

In his Doctrine of Virtue attempt to fill this gap Kant does not, in any case, rely on the notion of a commitment to represent our rational nature externally. But he does invoke the idea that in lying an agent treats himself merely as a means. Close examination of this notion therefore makes sense. It might not render Kant's derivations acceptable, but it should make them a bit more comprehensible.

III.

Kant writes little about precisely what treating a person merely as a means amounts to. But he suggests possible interpretations in his attempt in the *Groundwork* to show how a duty not to make false promises stems from the Formula of Humanity. As we noted, according to one influential interpretation Kant there implies that an agent treats another merely as a means and thus wrongly if in his treatment of the other the agent does something to which the other cannot consent. Yet we cannot effectively adapt this account to the case of an agent treating herself in some way. Since the agent is free, it is always possible for her to consent to the way she treats herself.

In the false promising passage Kant appeals not only to the idea that another cannot consent to the way an agent is treating her, but also to the idea that the other "cannot contain the end" the agent is pursuing:

> He who has it in mind to make a false promise to others sees at once that he wants to make use of another human being *merely as a means*, without the other at the same time containing in himself the end. For, he whom I want to use for my purposes by such a promise cannot possibly agree to my way of behaving toward him, and so himself contain the end of this action [*G* 4: 429–430].

Kant intimates that if another cannot contain or, less awkwardly, have the end an agent has in treating her in some way, then the agent treats the other merely as a means. Two agents presumably have a particular end if

the following is the case: they are both trying, or at least have both chosen to try, to realize this end. If this is not the case, then they presumably do not each have the end. But what, precisely, does it mean to say that an agent *cannot* have the end another possesses? Returning to the example at hand, what does it mean to say that the promisee cannot himself contain the promisor's end? From the outset it is important to specify precisely which of the promisor's ends the promisee cannot have. It is presumably the promisor's end of getting money from the promisee without ever paying it back. For the promisor's *ultimate* end might be that of diminishing child mortality, and there seems to be no reason why the two cannot both have that end. But it remains unclear just what sense of "cannot" Kant is invoking (or should invoke) in suggesting that a promisee cannot possess the false promisor's end.

Perhaps Kant's view is that the promisee cannot have the promisor's end in the sense that, in typical cases, it would be practically irrational for him to have this end.[12] (This is admittedly a conjecture. But the alternative to making such conjectures is to leave Kant's account

12 On a reading suggested by Thomas Hill, for Kant the promisee cannot share the promisor's end in that it is *logically impossible* for him to do so (Thomas E. Hill Jr., 2002, "Hypothetical Consent in Kantian Constructivism," in: *Human Welfare and Moral Worth*, Oxford: Clarendon Press, pp. 61–95, at pp. 69–70). Suppose the promisor, a borrower, has the end of getting money from the promisee, a lender, without ever paying it back. The borrower makes a false promise in order to secure that end. At the time he makes a loan on the basis of this promise, the lender cannot himself share the end of the borrower's getting the money from him without ever paying it back, goes this reading. If the lender shared the borrower's end, then he would not really be making a loan. For according to our practice, it belongs to the very concept of making a loan, as opposed, say, to giving money away, that one believe that what one disburses will be repaid. Given the aim of arriving at a plausible general account of treating others merely as means, this interpretation of the promisee's inability to share the promisor's end is unhelpful, or so I argue in detail in "Treating Others Merely as Means." For one thing, the interpretation renders uninteresting any generalization from the promising case. Suppose for a moment that a sufficient condition for an agent's treating another merely as a means is that it is logically impossible for the other to share the end the agent is pursuing in treating her in this way. Paradigmatic cases of treating another merely as a means do not involve such logical impossibility. Take, for example, a loiterer who threatens an innocent passerby with a gun in order to get $100. The sort of sufficient condition for treating another merely as a means that we seek should allow us to conclude that the loiterer is treating the passerby merely as a means; for he is mugging her. But the sufficient condition on the table does not do this. It is improbable, but still logically possible, that the passerby shares the loiterer's end of his getting $100.

extremely vague). In typical cases, it would be irrational for the promisee to try to realize the end of making a loan that is never repaid. For this end's being brought about would prevent him from attaining other ends he is pursuing, ends such as buying new rose bushes, saving money for retirement, and, of course, just plain getting his money back. The notion of irrationality at work here is familiar. In the *Groundwork*, Kant seems to embrace what Hill calls "the hypothetical imperative"[13], namely a principle that goes roughly like this: If you will an end, then will the means to it that are necessary and in your power, or abandon the end. Kant implies that the hypothetical imperative is a principle of reason: all of us are rationally compelled to abide by it.[14] An agent would act contrary to the hypothetical imperative and thus irrationally by willing an end yet, at the same time, willing another end, the attainment of which would, he is aware, make it impossible for him to take the otherwise available means to his original end. An agent would violate the hypothetical imperative, for example, by willing to buy a car yet, at the same time, willing to use the money he reserved for the down payment to make a gift to his nephew. The Kantian hypothetical imperative implies that it is irrational to will to be thwarted in attaining ends that one is pursuing. In typical cases, if a promisee willed the end of a false promisor, she would be doing just that.

Against the background of this example, we might interpret Kant to be claiming the following: If another cannot have the end an agent is pursuing in treating her in some way, then the agent treats the other merely as a means. The other cannot have the agent's end when the other cannot pursue it without practical irrationality of the kind we have just described. A sufficient condition for the moral impermissibility of using another is that it be done to attain an end that the other cannot have.

Are we able to adapt this account to cases of an agent's treating himself in some way? At the outset we might claim that if an agent cannot (rationally speaking) have the end he is pursuing in treating himself in some way, then the agent treats himself merely as a means. An agent cannot have this end just in case he cannot pursue it without practical irrationality – in particular without willing to be thwarted in his pursuit of some other end he has.

---

13 Thomas E. Hill Jr., 1992, "The Hypothetical Imperative," in: *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca: Cornell University Press, pp. 17–37.

14 For discussion in the *Groundwork*, see [*G* 4: 413–418].

Unfortunately, this account suffers from a serious problem. Suppose that a very rich person has the end of being the world's richest. But after taking a long trip to Africa, he decides to donate anonymously a large sum to a famine relief organization. He realizes that making the donation will prevent him from becoming the world's richest. Nevertheless, he refuses to give up his end. This person is undoubtedly behaving irrationally. But the account implies implausibly that his behavior is also morally wrong. For, according to it, he is treating himself merely as a means. In his giving away his money, he is willing to be thwarted in attaining his end of being the richest person on the planet. It would be easy to multiply cases such as this. The difficulty arises because in pursuing ends that are, intuitively speaking, very worthy, an agent might be thwarting his attainment of neutral or even bad ends.

In order to avoid this difficulty, we need to modify the account. I suggest the following: If an agent cannot have the end he is pursuing in treating himself in some way, then the agent treats himself merely as a means. An agent cannot have this end just in case – and here's the modification – he cannot pursue it without thwarting the pursuit of some end that he is *rationally compelled* to have. Of course, according to Kant our agent is not rationally compelled to have the end of being the richest person on earth. So he would not be treating himself merely as a means in donating a large sum to famine relief. In my view, this account coheres well with how Kant employs the notion of treating oneself merely as a means in his derivations of perfect duties to oneself. In the next section we will examine three of these derivations. There we will find some examples of ends that, in Kant's view, we are rationally compelled to have.

IV.

Let us begin with Kant's derivation of a duty not to kill oneself. Kant appeals to the notion that in killing oneself, one would be treating oneself merely as a means:

> To annihilate the subject of morality in one's own person is to root out the existence of morality itself from the world, as far as one can, even though morality is an end in itself. Consequently, disposing of oneself as a mere means to some discretionary end is debasing humanity in one's person (*homo noumenon*), to which the human being (*homo phaenomenon*) was nevertheless entrusted for preservation [*MM* 6: 423].

Morality is an end in itself, and without humanity, there would be no morality, Kant implies; for humanity is the "subject" of morality. Kant obviously also holds that humanity is an end in itself (e.g., [*MM* 6: 434–435], [*G* 4: 435]). Since it is, we are rationally compelled to view it as something "which must never be acted against" [*G* 4: 437]. Refraining from destroying humanity must be one of our ends. Now suppose that an agent is in pain, with no prospect of its abating as long as he lives. He has an end, namely that his suffering stop. Kant would call this a "discretionary end," which, I take it, is an end that an agent is not rationally compelled to have. In pursuing the end by killing himself, the agent would obviously be treating himself in some way. According to our interpretation, a person treats himself merely as a means if he does something to himself in pursuit of an end that he cannot himself have. That a person cannot have an end does not on this interpretation entail that it is impossible for him to adopt it, of course. A person cannot have an end, rather, if his willing it would be practically irrational in the sense of thwarting his attainment of some other end that he is rationally compelled to have. But in this case his willing the end that his suffering abate would clearly be practically irrational in this sense. In taking the means to his end, that is, in killing himself, the agent would render himself unable to attain an end that he is rationally compelled to have, namely that of refraining from destroying humanity. So it is not hard to see why, according to Kant, the agent would be treating himself merely as a means.

It is much more difficult to discern how, in performing certain sexual acts an agent would be treating himself in this way. Kant suggests that a desire to masturbate and thus masturbation itself are unnatural: "Lust is called *unnatural* if one is aroused to it not by a real object but by his imagining it, so that he himself creates one, contrapurposively; for in this way imagination brings forth a desire contrary to nature's end …" [*MM* 6: 425]. "Nature's end" refers to the preservation of the species. From a contemporary perspective, Kant's reasoning here is puzzling, even if we grant the legitimacy of thinking in terms of nature's having an end. For a given end is presumably contrary to nature's end only if the former's realization prevents, or diminishes the likelihood of, the latter's realization. But to many of us it seems odd to think that masturbation interferes with reproduction.

In Kant's time, however, it was a common view. In 1759, the well-respected French physician S. A. D. Tissot published *Onanism*, an 18th

century best-seller that soon appeared in German.[15] Tissot offers an extensive list of maladies that in his view stem from masturbation. They include pimples, weakness, gastrointestinal distress, shortness of breath, and loss of memory (Tissot, 1985: 26–27). Tissot also writes of "the indifference which this infamous practice leaves for the lawful pleasures of Hymen, even when their inclinations and powers still remain; an indifference which does not only induce many to embrace a life of celibacy, but even accompanies the nuptial bed" (Tissot, 1985: 43). Tissot was far from alone in decrying the ills of masturbation.[16] According to medical authorities contemporary with Kant, even if the practitioner of masturbation is physically able to reproduce, which is not a given in light of the many maladies that stem from the activity, he or she might simply lack any desire to do so.

In any case, Kant tells us that it is "not so easy to produce a rational proof" that masturbation is a violation of a duty to oneself. He then says that the "*ground of proof* is … that by it the human being surrenders his personality (throwing it away), since he uses himself merely as a means to satisfy an animal impulse" [*MM* 6: 425]. In masturbating, an agent does seem to use himself as a means to sensual satisfaction. But why does Kant think he uses himself *merely* as a means? In bending down to smell a rose, an agent also uses himself (e.g., his olfactory capacity) as a means to sensual satisfaction. Yet this agent does not presumably use himself merely as a means. Why not?

Kant does not offer an explicit answer to this sort of question. Nevertheless, against the background of the views on masturbation prevalent at Kant's time, an initial answer suggests itself. In masturbating, an agent undoubtedly treats himself as a means to gratify a certain animal impulse. He treats himself *merely* as a means for the following reason: in willing in this way to satisfy the animal impulse he destroys or diminishes

15 S. A. D. Tissot, 1985 [1766], *Onanism*, transl. by A. Hume, New York: Garland. For a discussion of the importance of Tissot's work, see Thomas W. Laqueur, 2003, *Solitary Sex: A Cultural History of Masturbation*, New York: Zone Books, pp. 37–40.

16 For discussion see Laqueur, 2003, pp. 25–63. Laqueur points out (60) that in 1786 the *Berliner Monatsschrift*, which had earlier published Kant's prize competition essay "What is Enlightenment?", solicited for another competition papers on the topic "How children and young people can be spared the physically and spiritually devastating vices of unchastity in general and onanism in particular, or, insofar as they are already infected by these vices, how they can be healed."

his ability to reproduce, according to 18th century experts. But since the reproduction of the human species is "nature's end," each individual is rationally compelled to have this end, continues this answer. So the agent behaves in a practically irrational manner. In typical cases, an agent's using himself in order to smell a rose is not practically irrational; in using himself in this way he would not be willing to be thwarted in the pursuit of any end that he is rationally compelled to embrace.

Kant does not make this argument, to my knowledge. In any case, it is implausible to claim that everyone, including a postmenopausal woman or an infertile man, is rationally compelled to have the end of reproducing. This claim seems to run afoul of Kant's dictum that ought implies can (e.g., [*CPrR* 5: 159]). But Kant seems to reject even the view that every fertile person must have the end of reproducing: "The end of begetting and bringing up children may be an end of nature, for which it implanted the inclinations of the sexes for each other; but it is not *requisite* for human beings who marry to make this their end in order for their union to be compatible with rights, for otherwise marriage would be dissolved when procreation ceases" [*MM* 6: 278]. Granted, that failure to have the end of procreation is compatible with rights does not itself entail that it is compatible with morality. But in the Doctrine of Virtue, Kant does not set out procreation as an imperfect, let alone as a perfect, duty.

An initial answer to the question of why, according to Kant, masturbation amounts to treating oneself merely as a means, whereas smelling a rose does not, has turned out to be unsatisfactory. But this answer suggests a related one that might capture at least part of Kant's thinking.

I mentioned that leading physicians of Kant's time believed that masturbation had deleterious effects on health. According to Tissot (1985: 75–76), one of these effects was an inability to focus on any project:

> A great number of young people are hereby greatly prejudiced, even when their faculties are not entirely destroyed, by their use being prevented. In whatever vocation a person is engaged, some degree of attention is required, which this pernicious practice renders him incapable of… I could enumerate those, whom this incapacity of fixing to any particular thing, joined to the decay of the faculties, had incapacitated to make a decent appearance in society. Shocking fate! which places man beneath the brute creation.[17]

17 Kant says that unnatural vice "debases [a human being] beneath the beasts" [*MM* 6: 425].

Perhaps, in accordance with Tissot, Kant believed that masturbation brought about significant impairment of rational capacities.

If so, we can with the help of our procedural account pinpoint the sense in which an agent who engages in masturbation treats himself merely as a means, according to Kant. As we found in our discussion of suicide, we are, according to him, rationally compelled to have humanity's preservation as an end. But in using himself (masturbating) in order to gain sensual pleasure, an agent would, in effect, be willing to be thwarted in preserving his humanity, that is, his rational nature. For he would be damaging his capacity to pursue ends – a capacity that is central to his rational nature. By contrast, in using his nose as a means to get pleasure from a rose, an agent would not be willing to be thwarted in preserving his humanity. So he would not be treating himself *merely* as a means.

On the interpretation just offered, an agent treats himself merely as a means in masturbating in essentially the same way as he does in committing suicide. He is rationally required to have the end of preserving his rational nature. In both cases, in treating himself in some way in order to secure pleasure or avoid pain, he in effect wills to be hindered in preserving it. But Kant holds masturbation to be even lower than suicide [*MM* 6: 425]. According to him, "someone who defiantly casts off life as a burden is at least not making a feeble surrender to animal impulse in throwing himself away: murdering oneself requires courage…" [*MM* 6: 425]. Kant seems to hold that engaging in masturbation amounts simply to succumbing to inclination. Someone who commits suicide might also succumb to inclination, for example, that to avoid suffering. But she must overcome other inclinations, for example, that to go on living, Kant suggests. So the agent who commits suicide manifests a strength of character lacking in the one who gratifies her "unnatural" lust.

Let us turn now to a third perfect duty to oneself, namely that not to lie. In telling a lie an agent disrespects his humanity, according to Kant. He does so, Kant suggests, by treating himself merely as a means. According to our interpretation, a person treats himself merely as a means if he does something to himself in pursuit of an end such that his doing it thwarts his attainment of something that he is rationally compelled to will. Is there an end that is both such that a person is rationally compelled to it and his lying would prevent him from realizing it?

Kant does not explicitly identify such an end. But he does refer to the "natural purposiveness of the speaker's capacity to communicate his thoughts" [*MM* 6: 429]. He also says that "the human being as a moral

being (*homo noumenon*) cannot use himself as a natural being (*homo phaenomenon*) as a mere means (a speaking machine), as if his natural being were not bound to the inner end (of communicating thoughts)" [*MM* 6: 430]. Kant seems to imply in the latter passage that since an agent as *homo phaenomenon* is bound to the end of communicating his thoughts accurately, the agent as *homo noumenon* is rationally compelled also to adopt this end. In short, he holds that an end everyone is rationally compelled to have in communicating his thoughts is to do so accurately [*MM* 6: 430].

It is now easy to see why, according to Kant, a liar treats himself merely as a means. He cannot pursue the end of the lie's taking place. For if he does he prevents himself from attaining an end that he is rationally compelled to have, namely that of communicating his thoughts accurately.

V.

Kant employs the notion of using oneself merely as a means in his derivations of three perfect duties to oneself. We have explored the content of this largely neglected notion. On our interpretation, if an agent cannot, rationally speaking, have the end he is pursuing in treating himself in some way, then the agent treats himself merely as a means. An agent cannot have this end just in case he cannot pursue it without practical irrationality – in particular without willing to be thwarted in his pursuit of some end he is rationally compelled to have. This account of treating oneself merely as a means has enabled us to grasp how from the Formula of Humanity Kant derives duties not to kill oneself, not to masturbate, and not to lie. Of course, it is one thing to understand these derivations, but quite another to judge them to be sound. All three raise significant philosophical issues.

The derivations of the duties not to lie and not to defile oneself with lust are open to obvious criticisms. The former rests on the assumption that whenever an agent communicates his thoughts, he is rationally compelled to have as an end to do so accurately. Kant suggests that this assumption is grounded in the "natural purposiveness of the speaker's capacity to communicate his thoughts". Yet Kant offers no justification for the notion that the natural end of this capacity is to reflect faithfully the contents of one's mind. Why couldn't its natural end, assuming it has one, be something else, say, that of promoting the individual's survival? In

that case, lying to preserve one's own life would of course not be contrary to the capacity's natural end. Kant's derivation, as interpreted above, of a duty not to masturbate has an even more obvious flaw. It rests on the long since discredited view that masturbation damages an agent's cognitive capacities. 18th century medicine appears not to have served Kant well.

Kant's failure to demonstrate that each of us has a duty to himself never to lie and never to masturbate stems from flaws in how he applies the Formula of Humanity, not necessarily from any shortcoming of the principle itself. Kant says that: "… a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular *nature* of human beings, which is cognized only by experience, in order to *show* in it what can be inferred from universal moral principles" [*MM* 6: 217]. In the two cases at issue, Kant relies on a questionable understanding of the particular nature of human beings in his quest to determine the implications for us of a universal moral principle.

His argument for a duty not to commit suicide raises a different set of difficulties, one of which I will sketch, without venturing to resolve it. The argument, at least as we have reconstructed it, relies on the premise that humanity must never be "acted against." If anything counts as acting against humanity, then destroying it does, it seems. So the premise appears to entail that an agent must never destroy humanity. But what about a case of self-defense? A powerful stranger attacks you suddenly with a knife. With the thought "one of us is going to die and better him than me," you turn your attacker's knife on him and kill him. You have acted against the attacker's humanity, it seems. Yet Kant hints that in his view you have not acted wrongly (see [*MM* 6: 235]). Or what about an executioner carrying out a death sentence on a convicted murderer? He appears to act against the murderer. Yet Kant makes it plain that in his view capital punishment for murder is justified [*MM* 6: 334].

One way to respond to these examples is to insist that, despite appearances, neither the executioner nor the victim of attack is really acting against anyone's humanity. An agent counts as acting against humanity only if he acts against a being that has dignity. But both the murderer and the attacker have, through their morally impermissible actions, forfeited their dignity.

On occasion Kant does suggest that dignity is alienable. "By a lie," he says, "a human being throws away and, as it were, annihilates his dignity as a human being. A human being who does not himself believe what he tells another … has even less worth than if he were a mere thing; for a

thing, because it is something real and given, has the property of being serviceable so that another can put it to some use" [*MM* 6: 429]. At least at the moment she lies, an agent jettisons her dignity, Kant here seems to imply. He also tells us that by unnatural use of his sexual attribute, a human being "surrenders his personality (throwing it away)" [*MM* 6: 425]. So at least while "defiling" himself with lust, an agent seems to forfeit his special value. If lying or masturbating can result in a loss of dignity, then it seems reasonable to hypothesize that committing murder or attempted murder can as well.

But the passages in which Kant seems to imply that dignity is alienable do not represent his considered view, in my opinion. For he claims repeatedly that humanity has an unconditional value that is beyond any price (e.g., [*G* 4: 434–435], [*MM* 6: 434–435]), at one point saying flatly that "humanity itself is a dignity" [*MM* 6: 462]. If the value of humanity is unconditional, then it has this value in every possible context in which it exists, including those in which the person who possesses it lies or masturbates. I suspect that in his discussions of persons who do these things Kant engages in a bit of exaggeration. They do not throw away their humanity, but rather, through disrespecting it, treat themselves *as if* they had none.

If this is correct, then Kant faces a challenge. He needs a principled justification for the claim that while an agent acts against himself in committing suicide in order to relieve his suffering, the crime victim and executioner in our examples do not act against anyone. And this justification should not be based on the notion that while the sufferer has dignity, the murderer and attacker do not. Of the three attempts Kant makes to derive a duty to oneself with the help of the notion that it is wrong to treat oneself merely as a means, the most straightforward is that to derive a duty not to kill oneself. But this attempt raises difficult issues concerning the relations between Kant's claims regarding the value of humanity and his claims regarding what we are morally permitted to do to it.

Duties to Others: Demands and Limits

Katja Maria Vogt

Reason prescribes, according to Kant, moral demands of different kinds. Taken together, these demands form a *system of duties* [*MM* 6: 242].[1] Duties to others need to be located within this system. As duties of virtue, or ethical duties, duties to others relate to duties of right, and within ethics, to duties to oneself.[2] Duties to others have, in the past decades, mostly been discussed with a view to Kant's conception of imperfect or wide duties, and their "latitude". If we want to understand the demandingness of duties to others, it seems we have to push for *precision* with respect to the notion of latitude. In this paper, I am suggesting an alternative approach, starting out from Kant's conception of a system of duties. Other duties, as it were, delineate the "space" of duties to others within a virtuous life. This suggestion may seem to imply that Kant's moral philosophy is highly demanding. If it is only duties – not ends – which limit other duties, life seems to be thoroughly structured by, and pervaded by duty. However, I will argue that for Kant, the duty to cultivate one's non-moral capacities significantly limits duties to others, and that his views on this duty capture a range of intuitions about how morality (understood as what we are to do *for others*) should leave room for our own lives. But still, according to Kant the agent's own concerns, her talents, and aspirations do not limit duties to others *as ends*, but *as duties*. Does Kant, at any point, allow the agent's ends to limit her duty? Kant is often quoted as referring to the agent's "true needs", and this is taken to indicate that he recognizes how the agent's *ends*, not her duties, delineate what we are to do for others. But as I will argue, Kant's remarks

1 Bracketed numbers refer to the volume and pages in the Preussische Akademie edition.

2 Throughout this paper, I am using "duties to oneself" and "duties to others" as referring to duties of *virtue* (in accordance with Kant in the *Doctrine of Virtue*), even though, in a sense, duties of right are also to oneself and to others (cf. [*MM* 6: 240 and 236]).

on true needs remain tentative, and are far less integrated into his ethics than one might wish.[3]

In contemporary moral philosophy, a wide range of views acknowledge some kind of limit to the demands of morality. This may be most obvious when the notion of supererogation is introduced.[4] Supererogationists, as we may call them, argue that there are actions that, while particularly laudable, are not required. But a theory need not focus on such a notion in order to acknowledge limits of moral demands. In the past decades, a number of ethicists have argued that agents have projects and close relationships which are so important to them, and indeed central to who they are, that morality should be understood in such a way as to leave room for them. Following Kagan[5] and Scheffler[6], I will refer to those who endorse such claims as "moderates".

---

3 My discussion of duties to others does not extend to Kant's justification of these duties. The demandingness of Kantian duties to others has, in the past decades, received considerable debate, and it is impossible to do justice to the subtlety of the discussion in a brief paper. The following publications have been particularly influential: Thomas E. Hill, Jr., 1971, "Kant on Imperfect Duty and Supererogation", in: *Kant-Studien* 62, pp. 55–76; Hill, 2002, *Human Welfare and Moral Worth. Kantian Perspectives*, Oxford: Clarendon Press; Marcia Baron, 1995, *Kantian ethics almost without apology*, Ithaca, NY: Cornell University Press; Nancy Sherman, 1997, *Making a necessity of virtue*, Cambridge: Cambridge University Press.

4 Discussion of this category has been importantly shaped by: Paul Eisenberg, 1966, "From the Forbidden to the Supererogatory: The Basic Ethical Categories in Kant's *Tugendlehre*", in: *American Philosophical Quarterly* 3, pp. 255–69; J. O. Urmson, 1958, *Saints and Heros*, in: A.I. Melden (ed.) *Essays in Moral Philosophy*, Seattle, pp. 198–216; Frances Myrna Kamm, 1991, "Supererogation and Obligation", in: *American Philosophical Quarterly* 28, pp. 273–285; David Heyd, 1982, *Supererogation: Its Status in Ethical Theory,* Cambridge: Cambridge University Press.

5 Shally Kagan, 1989, *The Limits of Morality*, Oxford: Clarendon Press, New York: Oxford University Press. Kagan's discussion of moderate positions focuses on deontological restrictions as well as "options" (permission for the agent to take her own life and projects as particularly important).

6 Samuel Scheffler, 1992, *Human Morality*, New York: Oxford University Press. In his earlier book, *The Rejection of Consequentialism* (Oxford: Clarendon Press/ New York: Oxford University Press, 1982), Scheffler introduces the notion of agent-centered prerogatives. As David O. Brink (1994, "A Reasonable Morality", in: *Ethics*, vol. 104, pp. 593–619) points out in a review of Scheffler, the moderate theory that he discusses in *Human Morality* seems to have much in common with the argument in favor of agent-centered prerogatives that Scheffler presented earlier.

Supererogationists hold that (i) the demands of morality are limited, (ii) one can go beyond these limits, and (iii) it is especially admirable, praiseworthy or good to do so.[7] The fact that Kant distinguishes between what is meritorious and what is owed may seem to indicate that Kant accepts the idea that one can, in some sense, do more than is required. But Kant's theory of virtue does not allow for the specific mix of evaluation and deontological categorizing which marks the notion of the supererogatory, since, according to him, no action can be good without being done from duty, and thus every good action must be done as part of one's adherence to duty.[8]

But what about the intuitions of the moderate? While the moderate may share the first two assumptions of the supererogationist, she may be hesitant with respect to (iii). It is not clear whether going "beyond duty" impoverishes our lives in significant ways, and whether doing more than morality demands really is admirable. Thus, we might construe a position according to which (i) the demands of morality are limited, and (ii) it is possible to go beyond these limits, but not an ideal to do so. If we think of Kant's remarks about the phantastically virtuous person – the person who is so affected by doing good that she is in some kind of feverish state, afflicted with unhealthy enthusiasm [*MM* 6: 408–9][9] – it would seem that Kant himself held such a view; one can, as it seems, do "too much". On closer inspection, however, this person is not really "too virtuous"; she is suffering from sympathy like from an affliction, and will be left exhausted. The truly virtuous agent has a tranquil mind [*MM* 6: 409]. Thus, the phantastically virtuous person merely appears to be too virtuous, and in fact lacks virtue in the best sense. Similarly, the person who doesn't treat anything as morally indifferent, not even whether she eats fish or meat, does not go beyond virtue; she is getting it wrong [*MM*

7 Cf., e.g., Richard McCarty, 1989, "The Limits of Kantian Duty, and Beyond", *American Philosophical Quarterly*, vol. 26, pp. 43–52; Phillip Montague, 1989, "Acts, Agents und Supererogation", in: *American Philosophical Quarterly*, vol. 26, pp. 101–111; James S. Fishkin, 1982, *The Limits of Obligation*, New Haven: Yale University Press, p. 11.

8 Cf., e.g., Sherman, 1997, 332 f. For a detailed discussion of this point see Section 2.

9 Throughout this paper, I am quoting from Mary Gregor's translation of the *Doctrine of Virtue*, and, even where I do not quote the text directly, adopt much of her rendering of Kant's language. Immanuel Kant, 1996, *The Metaphysics of Morals*, trans. and ed. by Mary Gregor, with an Introduction by Roger J. Sullivan, Cambridge: Cambridge University Press.

6: 409].[10] Kant cannot therefore accept (ii). However, he may still hold some version of (i). Kant does not accept the idea that one could be too virtuous; virtue is an ideal. But he may hold that the demands of the various kinds of duties are in some sense limited.

To discuss the demandingness of Kantian duties to others is as much an exercise in piecing together a variety of aspects of his theory, trying to see how he might answer the questions that interest us, as in trying to understand the questions that interest him. Kant does not seem to share the contemporary worry that morality might turn out to be too burdensome.[11] But he does discuss the demands of duties to others in various respects, qualifying the attitude in which we are to adhere to them, how far we are to go in complying with them, etc. According to the interpretation that I am suggesting, Kant's concern is how we can understand the place of specific duties within an overall virtuous life, and the different structures of duties of right on the one hand, and ethical duties on the other. Kant starts out from tables on how we are to distinguish between different kinds of duties, which he presents in the introduction to the *Doctrine of Right* [*MM* 6: 240–2]. His various remarks on duties to others seem to envisage them as part of a landscape in which we are to locate and understand them. In Kant's system of duties, duties limit each other, and understanding how Kant conceives of one duty involves importantly studying how the "space" of this duty is delineated by other duties.

After a brief general sketch of Kant's conception of duties to others (Section I), I will examine how different duties relate to duties to others, and more specifically, to duties of beneficence. I will argue that while the duty to morally perfect oneself (Section II), duties of respect (Section III), as well as duties of right (Section IV) shape and structure duties to others, they ultimately make them more, or at least not less demanding. Duties to develop one's own *non*-moral capacities, on the other hand, limit the demands of duties to others in significant ways. It is only in a few brief remarks that Kant suggests that something *other* than duty may limit duty. The agent is closest to herself, and she knows her "true needs" best

10 Cf. Baron, 1995, p. 91.

11 The notion of the supererogatory may also seem to suggest something like a divide between ordinary agents and heroes or saints, and this kind of divide seems to be in conflict with core intuitions of an enlightened ethics. The idea that someone would see herself as an ordinary agent and, as it were, marvel at the sacrifices of some hero, seeing her own life on a wholly different plane, certainly does not fit into Kant's moral philosophy.

(Section 6). On the most straightforward interpretation of the *Doctrine of Virtue*, it is duties to cultivate one's abilities and talents which limit duties to others. However, the few comments which Kant makes on how the agent is closest to herself may seem, while – at least on my understanding – not fully integrated into Kant's argument, most directly related to contemporary concerns.

I. Limits and Demands of Duties to Others in Kant's Ethics

Ethical duties are those which cannot be externally enforced; the necessitation of duty must come from self-constraint, not external constraint [*MM* 6: 219, 379 f.].[12] However, at some points Kant writes as if the basic difference between ethical duties and duties of right were that ethical duties are *wide*.[13] He further complicates matters by going back and forth between calling them wide and imperfect in a way which does not make it entirely clear how the distinctions between wide and narrow on the one hand, and perfect and imperfect on the other hand relate.[14] He then starts the main text of the *Doctrine of Virtue* with a discussion of *perfect* duties to oneself – a highly unexpected turn, since the introduction leads the reader to believe that all of ethics is about imperfect, or wide duties: the duty to cultivate one's non-moral faculties, the duty to cultivate one's will, as well as the duties of love and respect for others.

Thomas E. Hill's paper *Kant on Imperfect Duty and Supererogation* (1971) has in many ways set the stage for the discussion of Kantian duties

12 When Kant sums up his discussion of ethical duties, he makes this his first point [*MM* 6: 410]. Cf. also [*MM* 6: 406–7], where Kant writes that the concept of freedom makes it necessary to divide duties into duties of outer freedom and inner freedom.

13 In Kant's summary account, this is the third point [*MM* 6: 410]. The second point is that ethical duties do not give laws for actions, but only for maxims of actions.

14 At [*MM* 6: 390] Kant states that the wider the duty is, the more imperfect is the agent's obligation to action; up to this point, Kant had talked of wide duties, but he now goes on to speak of imperfect duties. The most difficult point lies in a striking discrepancy between introduction and main text. In the introduction, it would seem that it is due to the imperfection (and/or *latitudo*) of ethical duties that they call for casuistical reflection; however, in the main text, perfect duties to oneself lead into such reflections. My brief sketch of the duties of virtue does not depend on a specific view on the difficult question of how we should think about differences between the introduction and the main text.

to others. Hill focuses on the latitude of imperfect and wide duties.[15] While he emphasizes that ethical duties prescribe maxims for action, not particular acts [*MM* 6: 388–9], he initiates debate on what "act principles" are implied by ethical duties. In his latest discussion of these matters, Hill argues that "noone disputes that imperfect duties imply act principles" of the form "Sometimes, to some extent, one ought…"; the dispute is, on his account, rather on what *more* they imply, i.e., how we are to spell out such act principles more precisely.[16] While I do not wish to be the only dissenter on what Hill describes as scholarly consent, I propose that there may be a fruitful alternative to trying to pin down the precise form that such act principles might take. I share the general sense that the notion of latitude is at the heart of Kant's conception of duties to others, and that any discussion of their demands and limits must figure out how we are to understand this latitude. However, it seems to me that, instead of trying to make Kant's discussion more precise by translating duties to others into act principles, we can study his emphasis on how wide duties relate to other duties.

When Kant introduces the notion of *latitudo*, he writes that wide duties allow for a range in how to comply with them [*MM* 6: 389–90]. He goes on to say that this play-room (*Spielraum*) is not to be confused with the permission to make exceptions. Rather, the agent is supposed to figure out ways of complying with *all* her wide duties. For example, love for one's parents and love for one's neighbor limit each other; the virtuous agent will choose ways of complying with one of these duties that will not preclude compliance with the other. While this passage is sometimes referred to as indicating that Kant's ethics is highly demanding, I think that the most significant point about it is that, according to Kant, duties *limit each other*. The passage leaves it open whether our duties to parents and neighbors are more or less demanding; but it emphasizes that the virtuous agent will structure her life so as to be able to comply with both.

Further, Kant is concerned with explaining the peculiar structure of duties which cannot prescribe actions, but only maxims for action; or to put it slightly differently, with the question how there might be duties which do not make it clear what is to be done, but are nevertheless just as much duties as are duties of right. Ethical duties seem comparatively fuzzier, and it seems that one will often have to neglect some in order to

15 *Latitudo* is the Latin term that Kant gives in parenthesis when he explains how ethical duties are wide [*MM* 6: 389–90].

16 Hill, 2002, 204.

comply with others, and perhaps even to go about one's daily life. But at the same time, they are still duties, and the fact that they cannot be brought under an external law doesn't taint the necessity of their requirements, or their normative force. The duties of virtue are duties in the full sense.

Questions about the demandingness of morality are, in contemporary ethics, quite generally conceived as dealing with duties *to others*, asking what or how much it is that we are required to do for others, alleviate their suffering, etc. But within the Kantian framework, the question how much morality demands arises just as much with respect to duties to oneself as with respect to duties to others. Ethical duties which are wide are duties of "commission". Ethics, other than the doctrine of rights, obliges us to adopt *objective ends* or *ends which are duties* [*MM* 6: 383]. One's own perfection, and the happiness of others [*MM* 6: 385–6] have to be adopted as objective ends and duties. One's own perfection is a twofold end: we are required to cultivate our faculties [*MM* 6: 387; 392], and to cultivate our will up to the purest virtuous disposition (*bis zur reinsten Tugendgesinnung*) [*MM* 6: 386–7]. To make the happiness of others one's end means making the *permitted* ends of others one's end. While it is up to the others to decide what to pursue as their ends, it is up to me to disregard those of their ends which I do not take to contribute to their happiness [*MM* 6: 388]. Thus, both one's own perfection and the happiness of others are duties which are ends, and one might wonder how far one is to go in one's efforts with respect to each of these duties. While it seems, from the point of view of recent debates, more worrisome to understand what the needs of others demand of the moral agent, within Kant's theory, it could seem equally difficult to see how much, and what, one is to do to better one's own adherence to duty, or to cultivate one's talents.

This difficulty leads, as Kant explains, into casuistry. According to the introduction, it is because ethical duties are wide or imperfect that we get into casuistry, i.e., the discussion of problematic cases, which "call upon judgment to decide how a maxim is to be applied" [*MM* 6: 411]. Kant offers "Casuistical Sections" which exemplify how we are to apply judgment if we want to comply with duties of virtue. These sections of text do not have the status of "science" or metaphysics. In Kant's terms, it would be "dogmatics" to present the kind of considerations which are to be discussed here as science. Further, such reflections must remain "fragmentary" [*MM* 6: 411]. The structure of the main part of the *Doctrine of Virtue* loosely corresponds to this claim: Kant goes back and

forth between the exposition of duties of virtue and sections on "Casuistical Questions".[17] Some of the most interesting remarks on the limits and demands of duties to others have the status of casuistry, not of science. In assessing Kantian duties to others, we will have to accept, as I suggest, that some of the most substantial ideas which Kant discusses remain tentative.

II. Virtue as Moral Perfection

Kant's duties to others can be considered as more or less demanding in two ways – with respect to how one is to act *from* duty in adhering to them, and with respect to the maxims for action which are called for. I will begin with the first question.

In Kant's terminology, ethics is defined as the study of the duties of virtue, or as the doctrine of virtue. When Kant offers the Latin version of the German term *Tugendlehre – doctrina officiorum virtutis*[18] [*MM* 6: 381] – it is made explicit that the doctrine of virtue is a doctrine of the *duties* of virtue. Insofar as what Kant discusses under the title of virtue really is a set of duties, his theory of virtue might appear quite fundamentally different from the idea of virtue as it is conceived within

17 However, while it may seem in the Introduction that only imperfect duties invite casuistry, the text starts with a long section on perfect duties to oneself, each of them receiving a section with casuistical questions. Kant does not adhere to a regular pattern. Most surprisingly, imperfect duties to oneself do not receive any casuistical discussion.

18 It would be a large project in itself to study the notion of duty in Kant as relating to the term *officium.* In the *Doctrine of Rights*, Kant refers to Cicero when he explains why he calls his moral philosophy a doctrine of duties instead of a doctrine of rights [*MM* 6: 239]. As I hope to argue elsewhere, it is Cicero's peculiar appropriation of the Stoic notion of *kathêkonta*, translated as *officia*, which allows him to speak of *officia* of virtue. Once we, like Kant, then proceed to translate *officium* as duty, we arrive at the idea of duties of virtue. On key passages in Cicero, and on the notion of *kathêkonta* as translated in Cicero, cf. John M. Cooper, 1999, "Eudaimonism, Nature, and 'Moral Duty' in Stoicism", in: J. M. Cooper, *Reason and Emotion*, Princeton: Princeton University Press, pp. 427–448; also published in: *Aristotle, Kant and the Stoics*, Engstrom/Whiting (eds.), Cambridge: Cambridge University Press 1996, pp. 261–284. Cf. also the first sentence of the *Doctrine of Virtue:* "In ancient times "ethics" signified the *doctrine of morals* (*doctrina moralis*) in general, which was also called the doctrine of duties." Clearly, in classical antiquity, ethics was *not* referred to as a doctrine of *duties.* The transition only works if something like Cicero's interpretative translation is adopted.

"virtue ethics", both in the contemporary and the ancient senses. However, with respect to the demandingness of virtue, Kant shares, to some extent, ancient intuitions: virtue is an ideal, and in important respects an ideal with respect to what might be called the motivational disposition of the soul.

While Kant speaks about the soul, he does not use the term "motivation". Kant's discussion is about the agent's heart as the source of her action [*MM* 6: 441]. He talks of her *Gesinnung*, which determines what she will allow to be incentives (*Triebfedern*) for her action – reverence for the law, or her inclinations.[19] Thus, the term "motivation" is foreign to Kant's own vocabulary, and I intend to use it merely in a provisional way, that does not imply any claims about Kant's moral psychology. I am using the term for the limited purpose of bringing out an important affinity between Kantian and ancient ethics. Virtue is, for Kant, and according to major positions of ancient ethics, an ideal with respect to the source of one's action. Even though virtue must be practical, it is in this sense about motivation. Kant defines virtue as "*fortitude* with respect to what opposes the moral disposition *within us*" [*MM* 6: 380], i.e., strength in adhering to duty in the face of one's inclinations, or as "the moral capacity to constrain oneself" [*MM* 6: 394]. Virtue must *result from* "considered, firm, and continually purified principles" [*MM* 6: 384–5]; but to be virtuous is, most centrally, to act *from* duty.[20]

Virtue is thus relevant to one's adherence to all duties, including the duties of right. Kant calls "act in conformity with duty *from* duty" a universal ethical command [*MM* 6: 391; cf. 392, 387].[21] The external

19 Kant often uses the phrase "to act *from*" (where "from" translates "*aus*" – *Handeln aus Neigung, aus Pflicht*).

20 This point is somewhat related to Christine Korsgaard (1996, "From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action", in: Engstrom/ Whiting (eds.): *Aristotle, Kant and the Stoics*, Cambridge: Cambridge University Press, pp. 203–236, at. pp. 204–5). However, my point is more preliminary than Korsgaard's. Independently from the details of Kant's claims about how one should be motivated, it is interesting that, while his doctrine of virtue may seem to be a doctrine of *duties* of virtue, and thus quite different from major ancient theories, it is on the most basic level about the ideal state of the soul or heart as the source of action – and thus closer related to ancient conceptions than it is often thought.

21 The interpretation that I am proposing is not committed to a view on the much debated question how acting from duty relates to other incentives. Cf., e.g.,

law can enforce the actions that duties of right call for, but it cannot enforce *reverence* for rights, or make reverence for the law the incentive for one's actions. If compliance with the duties of right springs from reverence for the law, it can be called virtuous or ethical [*MM* 6: 394].[22]

Virtue, understood as the ideal to act *from* duty, is limitlessly or maximally demanding. Doing one's duty *from* duty might, as Kant observes, look at first sight like a narrow obligation, but it isn't. A human being cannot "see into the depths of his own heart so as to be certain, even in a *single* action, of the purity of his moral intention and the sincerity of his disposition" [*MM* 6: 392]. To be motivated in a virtuous way is a matter of continuous striving, and one can never fully know whether one has achieved what one has strived for, even with respect to particular actions. The first command of all duties to oneself is to know oneself, in the sense of scrutinizing oneself [*MM* 6: 441]. Kant explains how one cannot adhere too much to duty: "For really to be *too virtuous* – that is, to be too attached to one's duty – would be almost equivalent to making a circle too round or a straight line too straight" [*MM* 6: 433].[23] Virtue is an ideal, and it is a duty to oneself to strive for this ideal – the duty to perfect one's virtuous disposition.

Once we fully appreciate this, we see why Kant cannot make sense of the key intuition of the supererogationist: that there are actions which are not required but are good or virtuous. No action can be called good if it is not done *from* duty, and one can only act from duty when one acts in compliance with a duty. With respect to the duties to others this means that whatever action of beneficence an agent will choose to perform, her action will be an act of duty. No matter how enormous her effort will be,

Barbara Herman, 1983, "Integrity and Impartiality", in: *The Monist*, vol. 66, pp. 233–250.

22 In this sense, virtue is tied to freedom. In his notes to [*MM* 6: 382], Kant remarks that a human being is the more free, the more he can be constrained by the mere representation of duty.

23 In talking about the straight line, Kant may adopt a Stoic image. However, it may seem odd that Kant would appropriate this image – in Stoic philosophy, it mostly serves to show how one is either virtuous or not (a stick is either bent or straight, it is bent even if only slightly bent). Kant clearly does not agree that there are no degrees in virtue. As Schneewind points out with reference to a passage from Herder, Kant even calls the Stoic ideal of perfect goodness "nonsense". J. B. Schneewind, 1996, "Kant and Stoic Ethics", in: *Aristotle, Kant and the Stoics*, Engstrom/Whiting (eds.), Cambridge: Cambridge University Press, pp. 285–301, at p. 292. An even more direct reference to the Stoic notion of virtue is [*MM* 6: 405], but again Kant does not make it explicit that he is quoting Stoic ideas.

and no matter how purely she will be motivated, her action will not go beyond duty.[24]

However, with respect to duties of right, Kant acknowledges actions which go beyond what can be enforced by *external law.* In the general introduction to the *Metaphysics of Morals*, Kant makes the following remark: to do more in the way of duty than one can be constrained to do by the law is meritorious; to do exactly what the law requires is what is owed; and to do less than what the law requires is morally culpable [*MM* 6: 227].[25] It is also in this general introduction that Kant explains how ethics has duties in common with right [*MM* 6: 220]: duties of right and duties of virtue can interrelate so that, to do more than what the external law requires is to adopt the perspective of virtue.[26] Thus, the notion of what is meritorious as it is introduced here does not locate the agent's action outside of duty. What is meritorious will still be done from duty. With respect to the doctrine of right it might be tempting to think that Kant could accomodate a category of the supererogatory. But this would be a mistake. To do more than can be *externally* enforced is merely to move from the sphere of right to the sphere of virtue (and that is, the *duties* of virtue), not from duty to the supererogatory.

The duty to cultivate a virtuous disposition is maximally demanding: no matter which duty we are observing, we are called upon to be progressively more purely motivated in our adherence to it. And we can only comply with the duty to moral self-perfection insofar as we adhere to other duties. The duty to perfect one's virtuous disposition is in this sense dependent on our adherence to other duties, duties which are more directly about action. Duties to others are, as it were, one of the fields in which progressive moral perfection can take place. But does the duty of

24 Of course, we might consider adopting a relaxed or revised notion of supererogation, and then work out a way in which Kant may allow for supererogatory acts. Hill argues that Kant could accommodate something like supererogation insofar as a *particular* act of beneficence may, while it is a duty to adopt the maxim to help others, not be required. He mentions the objection that there is a sense in which *no particular* act is required, since ethics gives, in Kant's terms, laws for maxims of action, not for action (2002, pp. 215–6). It seems to me that, for any particular act of beneficence, once the agent performs it, she performs it *as an act of adherence to her maxim, which it is a duty to adopt.* In this sense, her act seems to me to be performed as an act of duty.

25 On the notion of merit in Kant's ethics cf. Paul Guyer, 2000, *Kant on Freedom, Law, and Happiness*, Cambridge: Cambridge University Press, p. 323 f.

26 Reverence for right is already meritorious (as opposed to the conformity of one's actions with the duties of right) [*MM* 6: 390].

moral perfection delineate in any way *what* (or *how much*) we are required to do for others?

Baron argues that the duty to moral perfection "affects how we are to carry out the duty to promote others' happiness". The less sensitive agent may not choose actions of helping others that she might choose if she were more sensitive; "the full spirit of it [i.e. the duty to help others] is not brought out until that duty is shaped and 'stiffened' by the duty to improve oneself morally".[27]

While I can see what Baron means by saying that the duty to moral perfection *stiffens* all of the agent's adherence to duty, it is not clear to me how it would *shape* duties to others. If we distinguish Kant's claims regarding how we should be motivated, and how we should act (cf. [*MM* 6: 218]), we can see why the duty to moral perfection runs through all of an agent's adherence to duty, and is, as it were, its backbone. But how far one has succeeded in achieving one's moral self-perfection is not about how often or to what extent one is able to do something for others, so that, if one were not far advanced, one would do (and would have to do) fewer things. Someone might adhere to the duty of beneficence quite consistently, without their motivations ever being particularly pure. Of course, the agent who is far advanced with respect to her moral perfection is going to adhere to the duties to others. But the fact that an agent has not made much progress with respect to her motivation by no means absolves her from adhering to the duties to others. The extent to which an agent has already perfected her motivational disposition does not translate into what the practical duties of beneficence require: which actions ethical duties call for is, in this sense, quite independent from the duty of moral self-perfection.[28]

Ethical duties, it seems, cannot be limited by anything that Kant would consider "empirical"; and whatever state of perfection one may

27 Baron, 1995, p. 100.

28 A more detailed analysis of the duty to moral perfection would require engaging with Kant's distinction between the command "be holy", which refers to working on the purity of one's disposition to duty, and the command "be perfect", which refers to trying to fulfill all one's duties (progressing from one perfection to another) [*MM* 6: 446–7]. While this would, on my interpretation, not lead to any further clarification of duties to others, it would be interesting with respect to Kant's notion of virtue. As I hope to argue elsewhere, one of the aspects in which Kant engages more with ancient accounts of virtue than it is often supposed is that Kant seems to consider questions about the unity of virtue (i.e., about how a notion of virtue in the singular relates to the idea that there are several virtues).

have attained at a given point in one's life would count as an empirical fact. What we are required to do is determined by the law, and nothing else. Kant states that a human being must consider himself as capable of resisting and conquering the impulses of nature – "he must judge that he *can* do what the law tells him unconditionally that he *ought* to do" [*MM* 6: 380]. In this statement, and in more detail in *Religion within the Limits of Reason Alone*, Kant makes the following claim about the relationship of "ought" and "can": if we ought to do something, then we must be able to do it. Thus, "ought implies can" does not provide a limit to the demands of morality; rather, whatever morality demands, we must be able to do it and must see ourselves as being able to do it.[29] At the end of the introduction to the *Doctrine of Virtue*, Kant proposes three principles, the third of which emphasizes this very point:

> Ethical duties must not be determined in accordance with the capacity to fulfill the law that is ascribed to human beings; on the contrary, their moral capacity must be estimated by the law, which commands categorically, and so in accordance with our rational knowledge of what they ought to be in keeping with the idea of humanity, not in accordance with the empirical knowledge we have of them as they are [*MM* 6: 404–5].

Whatever limits to the demands of duty to others there may be, they can only be ascertained in the moral law itself, not in such empirical facts as where a specific agent is at a given point on her path to virtue.

III. What We Owe and What Is Meritorious: Duties of Respect

The distinction between what we owe and what is meritorious is introduced again at the beginning of the *Doctrine of Virtue*, this time specifically with respect to the duties of virtue. Again, this may suggest that, among the things that one can do in adhering to duties to others, some are immediately required while others are good to do, but not required. But once we see how Kant spells out the difference between what is owed and what is meritorious, it is clear that this distinction makes duties to others *more* rather than less demanding.

According to Kant, by doing what we owe to others we do not put them under any obligation, while by doing something meritorious, we do

29 Cf. also *Religion within the Limits of Reason Alone* (II.1.b (76)).

[*MM* 6: 448].[30] Both what we owe and what is meritorious, is to be done *as one's duty.* The distinction between what we owe and what is meritorious does not place any actions of help outside the realm of duty. For Kant, both the question of how much we owe to others and the question of how much we should meritoriously be doing for others, concern *duties* of virtue. Kant emphasizes the perspective of the addressee of the action: It emerges from Kant's discussion that ultimately, to do one's best with respect to duties to others involves *pretending* that what is meritorious is owed, so as to not put the addressee under the impression that she is under an obligation. Thus, Kant does *not* envisage an agent who would rather think of whatever could be done to help others as *not* owed, so as to be less burdened with strict requirements. Rather, he makes it part of duty to others to be prepared to see things as owed which are not in fact owed.

Kant ties the distinction between what is owed and what is meritorious to a discussion of love and respect for others. Love and respect are, in one sense, feelings which accompany the carrying out of duties to others. At the same time, however, Kant refers to them (as practical attitudes) as duties. The duty of respect seems only negative, insofar as it requires not exalting oneself above others [*MM* 6: 449]. But Kant's discussion makes it clear that this duty is actually quite complex. Not only are we to respect others; we are additionally to act so as to preserve their own respect for themselves. Kant arrives at the following picture: we are under the obligation to help someone poor, yet while this obligation may in fact be such that our action will be meritorious, we should endeavor to make the poor person feel as if our meritorious help is indeed owed or "but a slight service of love", so that we "spare him humiliation and maintain his respect for himself" [*MM* 6: 449; cf. 453, §31].

Mere benevolence, which is the practical side of love, doesn't "cost us anything". But benevolence *results* in beneficence [*MM* 6: 449]. Beneficence demands that we adopt the ends of others as our ends; beneficence is the duty which concerns alleviating deprivation or helping

30 Guyer argues that the performance of any particular act of benevolence is always more than is strictly owed (2000, p. 327). While I cannot go into the details of this question, it seems to me that there might be acts of benevolence which do not put the addressee of the action under an obligation (we might think of relatively small acts of kindness – they might not put the addressee of the action under an obligation, and they might seem to be required by the duty of beneficence).

others in need [*MM* 6: 452, §29].[31] The distinction between what is owed and what is meritorious again comes in: Kant's precise formulation of the duty of beneficence calls for promoting the happiness of others in need, "without hoping for something in return" [*MM* 6: 453, §30]. Not hoping for anything in return involves – independently of whether the action may in fact be meritorious or not – to not make the addressees of one's actions perceive them as meritorious, so that they will not feel under an obligation "which always humbles the other in his own eyes"; it is even better to practice beneficence in secrecy [*MM* 6: 453; §31].

The fact that Kant begins his discussion of duties to others with these considerations indicates that for Kant, thinking about what and how much one is to do for others is placed within a theory that starts out, in an important way, from the dignity of persons. As Kant writes in the introduction, that human beings are ends is the "supreme principle of virtue" [*MM* 6: 395]. Respect for others and self-esteem delineate how we are to understand practical duties to others. Love admonishes human beings to come closer to another, and respect calls for *keeping a distance* [*MM* 6: 449]. This makes Kant's ethics, in some sense, even more demanding. While helping others, there is a sense in which it is important to "leave them alone", to not disrupt the way in which they are masters over their own lives.[32] We are to do what is meritorious *as if* it were owed, to spare others the humiliation of feeling that they will have to pay us back or are obliged to us in any other way, and as a result see themselves as somehow dependent on us. The importance of self-respect calls for reflection about what is to be done for others that is not primarily quantitative – asking how much time or resources we are to "sacrifice". We are not only to help others, but are also expected to think about our help from the "receiving end", trying to find ways of helping which don't

31 Kant discusses three duties of love: beneficence, gratitude, and sympathy. The duty of gratitude captures, as it were, the flipside of Kant's thoughts about putting others under obligation; gratitude is a "sacred duty", reflecting that one appropriately appreciates any kindness one has received [*MM* 6: 454–5].

32 Marcia Baron (2002, "Love and Respect in the Doctrine of Virtue", in: Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*, pp. 391–407) suggests that duties of love and duties of respect are presented to an implausible extent as differing from each other, and opposed to each other. To her, the idea that respect "bids us to hold back from others" is "intuitively odd" (392). If we think about distance as leaving it up to others how they want to lead their lives, and what they consider as integral to their happiness, it would seem to me that there is no such oddness.

involve placing oneself above the addressee of one's action. Further, sympathy with others is, just as beneficence, a duty of love. We have a duty not to shield us from the sight of suffering, to actually go to the places where we can learn about it, and to cultivate a compassionate nature [*MM* 6: 457].

IV. Duties to Others and Perfect Duties

Duties of right are *perfect* duties, and – according to the main text of the *Doctrine of Virtue* – some duties to oneself are perfect, too. Duties of right delineate the space of compliance with ethical duties. When we choose between different ways of getting hold of money that we plan to give to charity, we are supposed to rule out robbery, but we may take money from our bank account. It is only the permitted ends of others which we are to make our own ends; if someone sees his happiness as depending on our assistance in a robbery, we are supposed to exclude this element of his perceived happiness from the happiness of others that is to be our end.[33] If it is due to the narrowness of duties of right that they take this kind of precedence, we should expect that perfect duties to oneself have a similar status. When deliberating about what one may do to help others, murdering oneself (i.e., suicide in the fullest sense of the term) is ruled out as a means of possible help, just as is robbery.[34]

But there is more to the relationship between duties of right and ethical duties to others than that certain types of acts are ruled out. Duties of right are, on the one hand, duties of individual persons; but the way in which duties of right are brought under external law shapes the political situation in which these persons live. It is this side of duties of right which contemporary ethical theorists have viewed as of enormous importance, and Kant seems to have expressed many of the ideas which have become central to the debate. While he ends with questions rather than any definitive conclusions, the link between political institutions and ethical duties which Kant points out in the casuistical section on beneficence seems worth quoting at length:

33 It is this claim, i.e., the acceptance of "deontological restrictions", which Kant may be said to share with the kind of position which I refer to as moderate.

34 However, even though they are not wide, perfect duties to oneself lead into casuistical reflections.

> If someone who exercises over another (a serf of his estate) the greater power permitted by the law of the land *robs* the other of his freedom to make himself happy in accordance with his own choices, can he, I say, consider himself the other's benefactor because he looks after him paternalistically in accordance with *his own* concepts of happiness? Or is not the injustice of depriving someone of his freedom something so contrary to duty of right as such that one who willingly consents to submit to this condition, counting on his master's beneficence, commits the greatest rejection of his own humanity, and that the master's utmost concern for him would not really be beneficence at all? Or could the merit of such beneficence be so great as to outweigh the right of human beings? – I cannot do good to anyone in accordance with *my* concepts of happiness (except to young children and the insane), thinking to benefit him by forcing a gift upon him; rather, I can benefit him only in accordance with *his* concepts of happiness.
>
> Having the resources to practice such beneficence as depends on the goods of fortune is, for the most part, a result of certain human beings being favored through the injustice of the government, which introduces an inequality of wealth that makes others need their beneficence. Under such circumstances, does a rich man's help to the needy, on which he so readily prides himself as something meritorious, really deserve to be called beneficence at all? [*MM* 6: 454]

The considerations which Kant introduces here are of diverse theoretical status. (i) That one cannot do good to anyone in accordance with one's own concepts of happiness is part of Kant's metaphysics of morals; (ii) similarly, that human beings have a right not to be submitted to having a master has the status of metaphysics. (iii) But that much of the need which we see in others originates from the injustice of the government and the resulting distribution of wealth is an empirical claim, and since this consideration plays a major role in Kant's reflections, they must, as a whole, be situated outside of metaphysics. With a view to contemporary discussions, it seems relevant to note that Kant does not make the outcome of the dispute depend on who will get the empirical questions right. Rather, since empirical questions are involved, he locates the whole dispute outside of moral philosophy proper.

Further, while Kant introduces the very link between duties of beneficence and institutions which many contemporary theorists see as central to the debate about the demandingness of morality, he spells it out differently.[35] Kant's worry is not that duties to help those in need seem

35 For those who might want to cite Kant as a predecessor of key arguments in the debate about justice (and, today, global justice), it may be slightly disappointing that in the very next casuistical section Kant raises a different kind of consideration: what if all morality were captured in duties of right, so that no

overwhelmingly demanding as long as we understand them as duties of individuals, and that we can as it were "cut down" what is called for in individual action when we see that it is institutional change which is in fact needed. Kant's question is whether someone may see himself as someone else's benefactor if he alleviates his suffering, but is at the same time part of the institution which causes the deprivation.[36] And his reasoning again extends to the perspective of the addressee of help. Since duty is not only about how one treats others, but also about one's attitude toward oneself, we need to ask whether those who *accept* situations which are against their right violate a duty to themselves, thereby "rejecting their humanity".

The link between virtue and the doctrine of right which Kant suggests does not limit the demands of beneficence. Rather, it is about how the helping agent is to see herself and the addressee of her action. In cases where both are part of an unjust institution, the agent is not only not acting in a meritorious way, but is, in fact, not even adhering to a duty *of virtue* by helping the other person. What she does is not really *helping* (or being beneficent) at all, but is rather more fundamentally determined by her role as "master" (or whatever her role within the unjust institution may be) than by the acts that she performs within this role. Questions about justice thus, like Kant's discussion of the duties of respect, refer us back to the overall framework of Kant's ethics, and his views regarding the dignity of persons. Neither the agent nor the addressee of the action should be in the situation in which they find themselves, since it conflicts with the dignity of persons. And whatever the agent may do, it is tainted by this fact; it cannot fully qualify as beneficence, which does not, of course, imply that she should not nevertheless do it.

V. Duties to Others and Duties to Cultivate One's Non-Moral Capacities

Duties to oneself limit and determine duties to others, and vice versa; in the agent's life, there must be room for both. And much of what *we* would describe as the agent's ends (her projects, plans, etc.) might be

room were left for beneficence as a duty? In this case, "a great moral adornment, benevolence (*Menschenliebe*), would be missing from the world" [*MM* 6: 458].

36 Kant also suggests that a rich person should not, when helping others, make his efforts appear meritorious, or regard them as such [*MM* 6: 453].

captured in the Kantian realm of duty: all those pursuits that belong to developing one's talents, one's intellectual and physical capacities, etc. Thus, many activities which may seem to us like the agent's "projects" – studying, getting educated so as to be able to hold a profession, learning to play the piano, exercising, etc. [*MM* 6: 444–5] – may fall, for Kant, into the sphere of duty to oneself.[37]

More broadly, what we might refer to as (in some sense) *values* or part of the agent's good life, is in Kant discussed in terms of *duty.*[38] Developing one's talents, for example, is not something that we value so highly that morality should make room for it; rather, agents have a duty to develop their talents. But like duties to others, duties to oneself are wide – noone can explore all talents she might possibly have, or know for certain whether the particular talents she decides to develop are in fact her greatest talents. Thus, duties to oneself leave the same kind of leeway that characterizes duties to others, and ethics as a system of ends calls for finding a way of life in which there is room for complying with all of them.

The duty to cultivate one's capabilities significantly delineates duties to others. Duties to help others may, for example, conflict with the duty to push forward with a research project. Neither the agent who devotes all her energy and time to her research, nor the agent who devotes all her resources to helping others is, on Kant's view, doing her duty. Contemporary ethicists tend to think only of one, broad kind of moral duty – duty to others. Thus, if they claim that an agent does more than is required, they will most likely view this as laudable. The space into which the agent's actions extend is a non-moral space; what the agent may neglect by doing more in the way of duty are her own ends, which are

37 While Susan Wolf's worry about the unattractiveness of leading the life of a moral saint does not disappear within such a framework, it would have to be rephrased. Contemporary concerns with a morality that "crowds out the non-moral virtues, as well as many of the interests and personal characteristics that we generally think contribute to a healthy, well-rounded, richly developed character" (Wolf, 1982, "Moral Saints", in: *The Journal of Philosophy*, vol. 79, no. 8, pp. 419–439, at p. 421) rest on a picture according to which morality is concerned fundamentally with what we are doing *for others.*

38 The question whether Kant can make room for what is important to the agent (issues which have been discussed largely by reference to the agent's identity and her integrity) has been raised to a great extent through engagement with Bernard Williams, and received much attention by Kant scholars and contemporary Kantians. I cannot here engage with this debate, which would – on my reading – call for discussion of the conception of duties to oneself.

considered as *morally* (though not otherwise) indifferent. But within Kant's framework, many of the things one would neglect have *moral* status. A life which does not make room for learning something, or for pursuing one's talents is not a life of praiseworthy sacrifice, but rather a life in which things of moral significance are being neglected.

What is more, duties to others and duties to oneself are connected. Each human being has the duty to be a useful member of the world, which one cannot be without developing one's talents. If one does not put effort into developing one's talents, one is, in the end, not really going to be able to act to the benefit of others. On the whole, an agent is more "useful", if she has developed her talents [*MM* 6: 445–6]. But this reference to one's "usefulness" does not lead Kant into moralizing one's choices with respect to education and career. Which talent one wants to develop is up to one's own reasonable reflection, in accordance "with his own rational reflection about what sort of life he would like to lead and whether he has the powers necessary for it (e.g., whether it should be trade, commerce, or a learned profession)" [*MM* 6: 445, §20]. A certain kind of pleasure in one's own life is part of duty to oneself: it is against duty to oneself to deny oneself the pleasure of some joys in life through exaggerated avarice or discipline [*MM* 6: 452].

Duties to oneself are, in Kant's ethics, the most straightforward limit to duties to others. As I have argued, the ways in which duties of respect and duties of right delineate duties to others make adherence to duty, in the end, more (or at least not less) demanding. Duties to others need to be complied with, as it were, under the premise of the duty of respect; concern for the self-esteem of those who are in need of help surely makes beneficence ultimately more demanding. Perfect duties delineate the range of choices that one has within the sphere of virtue. But more importantly, in situations where others are in need, we may have to consider whether we are part of the institutional framework which contributes to their deprivation, and may not be benefactors at all (I write "may" since this is, for Kant, an open, casuistical question). Thus, while the justice or injustice of political institutions shapes how we are to see our help for others, it does not make duties to others less demanding. But duties to oneself indeed make room for pursuing the life that one is interested in, the life in which one cultivates one's talents, has a fulfilling career, etc.

Recent scholarship often engages with Kant not in terms of interpretation or exegesis, but in terms of what *we*, were we to adopt a Kantian framework, could say about a certain question. From this point

of view, it would seem that Kant's most substantive limit of duties to others is closed off – hardly anyone today is convinced by Kant's arguments for duties to oneself. We might still find it plausible that one is to strive to become a better person. But the conception of a duty to cultivate one's non-moral capacities finds almost no adherents today. If we do not recognize such a duty, we are not rejecting a minor detail in Kant. Kant's moral philosophy is in a very fundamental way about a *system* of duties. Rejecting any of these duties will therefore be a substantial departure from this system. Thus, it seems that we can either, as interpreters, come to the view that duties to cultivate one's non-moral capacities are Kant's most significant limit to duties to others, or we can, as ethicists, recognize an important divide between a contemporary version of Kantian ethics that we might wish to hold on to, and Kant's ethics.

VI. Duties to Others and One's Own Ends

But Kant may have further arguments regarding the limits of duties to others, arguments which do not concern duties to oneself, but rather one's own ends. Such aspects of Kant's thought would be much closer to the concerns of contemporary ethics. I will conclude by examining a few passages which are as interesting as they are frustrating – Kant seems to acknowledge some of the intuitions that are important to moderates, but he does not spell them out.

In the introduction to the *Doctrine of Virtue*, Kant points out how the duties to oneself relate to one's own ends: an agent who gives away so much of her own means that she falls into poverty might endanger her own morality (*Sittlichkeit*) [*MM* 6: 388]. Thus, one's own happiness, while not a duty, is tied to the duties to oneself. One's moral perfection is – as Kant supposes – easier to attain in some circumstances than in others; and these circumstances happen to be those in which the agent is taking care of her own needs. Kant repeats this point in the main text, this time as a casuistical reflection [*MM* 6: 454]. It is important to note that the introduction does not distinguish between "science" and "casuistry", and we must therefore go to the main text to determine the theoretical status that Kant ascribes to particular considerations. That poverty makes human beings liable to vice is an empirical claim, with which one may or may not agree (we may for example think of an agent who is very certain that poverty will not mislead her into vice). It is

important to keep in mind what Kant himself thinks about the status of such empirical considerations. We cannot draw on the casuistical sections and claim that, in Kant's moral philosophy, we find such-and-such limits to the duties of others, or, as it is sometimes put, an indirect duty to promote our happiness. What we find is an empirical reflection on how our own virtue may relate to our happiness, and this reflection may play some role in how we are to adhere to the imperfect duties to others.

But what about the agent as "one of all"? Duties to others are universal duties. Strictly speaking, they are not duties to *others*, but to all human beings, of which the agent herself is one. Kant writes on the maxim of benevolence: "I want everyone else to be benevolent toward me (*benevolentiam*); hence I ought also to be benevolent toward everyone else. But since all *others* with the exception of myself would not be *all*, so that the maxim would not have within it the universality of law, which is still necessary for imposing obligation, the law making benevolence a duty will include myself, as an object of benevolence, in the command of practical reason". This doesn't make benevolence toward oneself a duty; but it does *permit* benevolence toward oneself "on the condition of your being benevolent to every other as well" [*MM* 6: 451]. Thus, a minimal criterion limiting benevolence toward others is benevolence toward oneself. However, as long as we think of this just insofar as the agent is one of all, this may mean very little. Any kind of significant permission to attend to one's own ends, it seems, would have to attribute some kind of special importance to them.

In the introduction, Kant mentions an additional aspect, a "maxim of promoting others' happiness at the sacrifice of one's own happiness, *one's true needs*, would conflict with itself if it were made a universal law" ([*MM* 6: 393]; emphasis K.V.).[39] How far we should go in the sacrifice of

39 In her assessment of the duty to promote the happiness of others, Andrea Marlen Esser refers to Kant's remarks about poverty as a temptation to violate one's duty and to this passage. She seems to find Kant's empirical claims about how need may endanger virtue quite compelling and, within Kant, unproblematic. On her reading, these passages once and again show that the image of Kant as "verknöcherte[r], lustfeindliche[r] Philister" is but a bad caricature. Kant, on her interpretation, *understood* that the good deeds of someone who neglects her own happiness will, in the long run, make the benefactor resentful, and thus be counterproductive (2004, *Eine Ethik für Endliche. Kants Tugendlehre in der Gegenwart*, Stuttgart-Bad Cannstatt: frommann-holzboog, pp. 341–342). Sherman thinks that this passage "reminds us that the Kantian agent's happiness and needs are not a matter of moral indifference" (1997, p. 339). However, as I will

our own welfare "depends, in large part, on what each person's true needs are in view of his sensibilities, and it must be left to each to decide this for himself" [*MM* 6: 393]. One might wish that Kant had elaborated more on this point; one's true needs, as understood by oneself, may indeed be a quite significant criterion for limiting what one is to do for others.[40] While Kant does not go on to explain the notion of one's true needs, there is a second passage which is similar in spirit, though it may seem to be even more difficult to interpret.[41] Kant discusses how the duty of benevolence relates to what he calls the ethical law of perfection – "love your neighbor as yourself":

> Yet one human being is closer to me than another, and in benevolence I am closest to myself. How does this fit in with the precept "love your *neighbor* (your fellowman) as yourself"? If one is closer to me than another (in the duty of benevolence) and I am therefore under obligation to greater benevolence to one than to the other but am admittedly closer to myself (even in accordance with duty) than to any other, then it would seem that I cannot, without contradicting myself, say that I ought to love every human being as myself, since the measure of self-love would allow for no difference in degree [*MM* 6: 451, §28].[42]

Kant talks about benevolence, not beneficence. But *practical* benevolence *is* beneficence [*MM* 6: 452]. Does Kant ask us to envisage something like circles of closeness, and a duty to take care of someone's needs in accordance with relative closeness, with the agent herself at the center of the circle? On this account, we are not only closest to ourselves by knowing our own needs best, but also *in duty.* This may seem to stand in stark contrast with Kant's thesis that there can be no duty to promote one's own happiness.

---

go on to argue, it is quite unclear how Kant would spell out the notion of one's true needs, properly understood only by oneself.

40 In his *Lectures on the History of Moral Philosophy* (2003, ed. by Barbara Herman, Cambridge/Mass. and London/England: Harvard University Press, p. 172 f. and p. 221 f.) Rawls emphasizes – albeit in a different context – that Kant's notion of human needs is not really spelled out, and that it is difficult to see how Kant would explain this passage.

41 Kant mentions one's true needs also in the context of duty to oneself: avarice, understood as "restricting *one's own* enjoyment of the means of good living so narrowly as to leave one's own true needs unsatisfied", is a vice which is contrary to one's duty to oneself as a moral being, i.e., a perfect duty to oneself [*MM* 6: 432]. Again, he does not spell out the notion of one's true needs.

42 In his earlier discussions of "Love your neighbor as yourself" in the *Groundwork* [*MM* 4: 399] and the *Doctrine of Virtue* [*MM* 6: 402], Kant had emphasized that the point of this command is to "do good".

Kant doesn't offer further argument, and the passage remains puzzling. Like no other passage in the *Doctrine of Virtue*, it recalls both a Stoic idea, and what Cicero makes of it. The Stoic theory of *oikeiôsis* presents the picture that Kant invokes – concentric circles of closeness with the agent at the center (however, not in Cicero's account of it).[43] But while the Stoic theory concerns the way in which one should try to develop the same disposition towards those from the outer circles as one has for those in the inner circles, Cicero himself does not adopt this idea.[44] According to Cicero, our *officia* are indeed more demanding with respect to those who are closer, and Kant may only spell out a further implication of this reasoning when he states that one is closest to oneself even in duty. A different (and extensive) study would be needed to find traces of Kant's engagement with Cicero (as a source and as a philosopher), and to assess it. However, at this point, one cannot help but feel that Kant presents a line of reasoning which is not fully integrated into his ethics. If it were, Kant would offer a very substantial criterion for limiting the demands of duties to others.

To conclude, let me briefly return to the notion of one's true needs. As Rawls emphasizes, we can only speculate about how Kant would explain this were we to push him for clarification.[45] Kant's claim that one is to judge one's true needs in view of one's own sensibilities seems to suggest something about how others are to keep a distance from the agent if they are to duly respect her. Kant writes that noone has the right to require of me that I sacrifice my ends if these are not immoral [*MM* 6:

43 Cf. Cicero, *De finibus* III.16 f. For a range of passages on the Stoic theory, cf. A. A. Long/D. N. Sedley, 1992, *The Hellenistic Philosophers*, ch. 57, Cambridge: Cambridge University Press.

44 Cf. Cicero's discussion in *De officiis* I, 15 (49) f. Cicero argues that there are several degrees of human *societas*, and the society of all human beings will, on his view, be best sustained if one dedicates more of one's well-doing (*benignitas*) to those with whom one is more closely connected. Cicero relates this idea to other considerations like need, or whether someone is dependent on the agent; however, for him it follows from the relative scarcity of resources that one should give priority to those who are close.

45 Rawls' engagement with this point is part of a different train of thought – Rawls is not pointing to its difficulty with a view to limits of the duties to others. He says: "I understand Kant to say that we have certain true human needs, certain requisite conditions, the fulfillment of which is necessary if human beings are to enjoy their lives." Rawls suggests that we may add to Kant (or amend his theory) by spelling out this notion in the sense of what humans *in fact* need. Ed. Herman (2003), p. 174.

388]. Along these lines, it might seem that, within Kant's moral philosophy, the distance that others are to keep from the agent reserves some kind of space for the agent to figure out what her most important ends are, what it is that she judges to be indispensible to her happiness, and to her enjoyment of life. If she errs, this is, as it were, the lesser of two evils. It might seem worse to interfere with the kind of space that respect calls for than to have someone go along with some misconception of her true needs.[46]

46 I am very grateful to Christian Barry for proofreading this paper, and for offering many helpful suggestions.

Virtue Ethics, Kantian Ethics, and the "One Thought Too Many" Objection

Marcia Baron

I.

Claims by proponents of virtue ethics that virtue "ethics is a rival to deontological and utilitarian approaches, as interestingly and challengingly different from either as they are from each other,"[1] have met with some skepticism. Kantians (and others) have wondered whether virtue ethics and Kantian ethics are as deeply opposed as some virtue ethicists claim, and if not, whether it might be possible to have a Kantian virtue ethics.[2] Such reflections were helpful in advancing the debate at a time when virtue ethicists did not provide a very clear picture of just what virtue ethics was, and often presented it primarily in terms of its opposition to Kantian ethics[3] and sometimes also to utilitarianism (or, more broadly, consequentialism). The challenges spurred virtue ethicists on to clarify and refine their conception of virtue ethics. In addition, the bar was raised for discussions of Kantian ethics, for as Kantians (myself among them) pointed out various ways in which Kantian ethics seemed to be compatible with virtue ethics[4] (seemed, because we couldn't tell

---

1 Rosalind Hursthouse, 1999, *On Virtue Ethics*, Oxford: Oxford University Press, p. 2.

2 The same issue can be raised concerning utilitarianism, but my focus here will be on Kantian ethics.

3 They sometimes spoke instead of "deontological ethics" or "deontology", apparently intending this to be equivalent to or to include Kantian ethics. I think "deontology" a rather unhelpful, because potentially misleading, term, so I'll stick to the term "Kantian ethics."

4 I did so primarily in my contribution to *Three Methods of Ethics* (Marcia Baron/Philip Pettit/Michael Slote (eds.), 1997, *Three Methods of Ethics: A Debate*, Oxford: Blackwell Publishers) and in a comment I presented on Rosalind Hursthouse's invited paper, "Applying Virtue Ethics," at the 1989 Pacific division meeting of the American Philosophical Association. I also explored the issue of just how virtue ethics and Kantian ethics differ in some earlier work (Marcia

exactly what virtue ethics was supposed to be), we corrected some misconceptions about Kantian ethics implicit in the characterizations of virtue ethics.

For some time the problem of ascertaining just what virtue ethics is supposed to be persisted, complicated by the diversity of views. To the extent that it was possible to discern what each proponent of virtue ethics took virtue ethics to be, the proponents seemed to have rather different ideas, though they were united in holding that virtue ethics was opposed to Kantian ethics, and was a viable alternative to the usual options.[5] An added complication was that others voiced similar criticisms of Kantian ethics and utilitarianism (or contemporary moral theory, or "Enlightenment morality") without using the label "virtue ethics." Should they too be considered virtue ethicists?

The possibility that virtue ethics and Kantian ethics are compatible was well worth exploring as long as contemporary virtue ethics remained very hard to pin down. An appropriate response was to lay out the various conceptions of virtue ethics and consider, on each conception, just what the possible variations are. We could then ask, with respect to each conception, whether a Kantian version is possible. If so, we could then conclude that Kantian ethics and virtue ethics were compatible after all, on at least one understanding of virtue ethics.

I believe that there is less value now in exploring that possibility. Thanks especially to the work of Christine Swanton, there is now considerably more clarity about virtue ethics, and therefore far less reason

Baron, 1985a, "The Ethics of Duty / Ethics of Virtue Debate and Its Relevance to Educational Theory", in: *Educational Theory*, vol. 35, no. 2, pp. 135–149), and laid out some ways of thinking about virtue ethics in Marcia Baron, 1985b, "Varieties of Ethics of Virtue", in: *American Philosophical Quarterly*, vol. 22, no. 1, pp. 47–53.

5 Some who place an emphasis on virtue and character (myself included) see such an emphasis as compatible with utilitarianism, Kantian ethics, or both. Partly because of this divide between those who think they are compatible and those who think they are not, it has more recently become common to understand "virtue ethics" more narrowly, so that it does not encompass just any approach that takes virtue and character very seriously, or even every approach that puts its emphasis on virtue or character (or virtue and character). Julia Driver, for instance, clarifies that her project is not (a form of) virtue ethics: "Virtue ethics is the project of basing ethics on virtue evaluation. I reject this approach. This is an essay in virtue theory, since what I am trying to do is give an account of what virtues are" (Julia Driver, 1996, "The Virtues and Human Nature", in: Roger Crisp (ed.), *How Should One Live?*, Oxford: Oxford University Press, pp. 111–129, note 1).

for Kantians to try to figure out just what virtue ethics is (and what it might be). Of course that might seem only to make the exploration of a possible Kantian virtue ethics easier, not to render it unnecessary or less worthwhile. But as I'll explain shortly, virtue ethics (as articulated by Swanton) and Kant's ethics seem pretty clearly incompatible; and this is no surprise, since the characterization of virtue ethics was crafted to rule out Kant's ethics.

Swanton distinguishes between virtue ethics as a genus and specific conceptions of virtue ethics, each of which is a species of the genus. There is thus an array of species of the genus virtue ethics, among them Neo-Aristotelianism and Swanton's pluralistic virtue ethics. With this in mind, we can make sense of there being a variety of types of virtue ethics and have no reason to ask which of the various proponents speaks for virtue ethics – provided, that is, that we can see them as linked together by being species of the same genus. What is the genus, of which Neo-Aristotelianism and Swanton's pluralistic virtue ethics are species? Swanton puts forward the following as a way of characterizing virtue ethics as a genus: "In virtue ethics, the notion of virtue is central in the sense that conceptions of rightness, conceptions of the good life, conceptions of 'the moral point of view' and the appropriate demandingness of morality, cannot be understood without a conception of relevant virtues."[6]

Swanton's approach is very helpful in that it permits the variation among virtue ethicists that has been evident in the contemporary virtue ethics literature, while at the same time providing an understanding of virtue ethics that unites them. It also leaves room for the development of new species of virtue ethics, or for arguing that a theory not generally seen to be compatible with virtue ethics (the genus) in fact is compatible with it (and perhaps even lends itself nicely to being developed as a version of virtue ethics, i.e., as another species of virtue ethics). Without taking Swanton's characterization of the genus to have canonical status, I think her approach (including the distinction between the genus and the

6 Christine Swanton, 2003, *Virtue Ethics: A Pluralistic View*, Oxford: Oxford University Press, p. 5. Swanton presents this characterization cautiously, prefacing it as follows: "'Virtue ethics' resists precise definition, and rightly so. For as I mentioned, it is frequently observed that virtue ethics in its modern development is still in its infancy. It should not therefore be shackled by preconceived ideas about its progeniture and nature."

species) the best way, at present, to understand virtue ethics, and in this paper I understand virtue ethics accordingly.

Another respect in which her characterization is helpful is that while fairly generously ecumenical, it is not a catch-all. It does not include under the heading of "virtue ethics" each and every approach that emphasizes character and virtue. That one offers an account of the virtues, or finds it fruitful to capture a lot of what in twentieth century ethics has been (purportedly) captured by talk of particular actions (or states of affairs) by focusing instead on character and qualities of character, is not enough to make one a virtue ethicist. I follow Swanton in using the term "virtue ethics" in a way that marks virtue ethicists off from those who place a considerable emphasis on virtue or offer an account (or theory) of virtue or the virtues, but are not committed to viewing virtue ethics as a distinct kind of normative theory (or, if one prefers not to use the term 'theory' here, approach). I do not take "virtue ethicist" to encompass everyone who thinks that virtue and character are neglected in many discussions of ethics and need to be given a more prominent place in ethics and ethical theory. Virtue ethicists put forward virtue ethics (whether as a theory or in an anti-theoretic way) as a rival to utilitarianism and Kantian ethics, and have sought to characterize it in a way that sealed that conclusion.

Keeping in mind the genus as Swanton characterizes it, I think it is safe to say that it would be an unwarranted stretch to view Kant's ethics as a species of virtue ethics. Not that it would be impossible to force it into that mold; we might argue that while Kant's conception of rightness can be understood without a conception of such virtues as gratitude [*MM* 6: 455–456][7] and modesty ("willing restriction of one's self-love in view of the self-love of others" [*MM* 6: 462]), it cannot be understood without a conception of the virtue of conscientiousness, or the virtue of being committed to perfect oneself. Or we might classify the good will as a virtue, and thereby ensure the conclusion that a conception of rightness cannot be understood, on Kant's ethics, without a conception of the relevant virtues (or at least a relevant virtue). The problem with such claims is not that they are false, but that they are forced. Kant does not present the good will as a virtue, and although it is less of a stretch to speak of the virtue of conscientiousness or the virtue of being committed

7 Page numbers refer to the Academy edition. The translations of Kant's used in this essay are those of Mary J. Gregor, in *Immanuel Kant: Practical Philosophy*, Cambridge: Cambridge University Press, 1996.

to perfect oneself, he does not present them in quite that way. Nor, more importantly, is it clear that a conception of rightness is impossible without a conception of conscientiousness or being committed to perfect oneself.

In thinking about whether it is valuable to try to merge Kantian ethics and virtue ethics, we should keep in mind that the category "virtue ethics" is one that has been developed quite recently, and for purposes that reflect current issues in contemporary ethics rather than for purposes that Kant had, or that relate helpfully to Kant's aims. Moreover, as noted, the focus has been on developing a conception of virtue ethics that allows it to be an alternative to – and incompatible with – both utilitarianism and Kantian ethics. This is particularly evident in the work of Michael Slote and Rosalind Hursthouse, and also informs Swanton's *Virtue Ethics.*[8]

But could we fight this trend – this trend of understanding virtue ethics as an alternative to, and at odds with Kantian ethics (and utilitarianism)? It might be possible. We could develop an account of various Kantian virtues (not necessarily enumerated by Kant), formulating them in such a way as to ensure that they meet Swanton's criterion. What I doubt is that this would be worthwhile, rather than simply combative. It would make sense only if we think that virtue ethics, as characterized by Swanton, would provide a shape for Kantian ethics that enhances it, and I see no reason to think that it would. In addition, I see no reason to resist the direction that Slote, Hursthouse and Swanton (among others) favor. If they think it valuable to develop various species of virtue ethics that take as their starting point that virtue ethics is a rival to Kantian ethics, I see no need to challenge them and to claim that in fact it is not really a rival to Kantian ethics, or that it need not be.[9]

---

8 See Hursthouse, 1999; Michael Slote, 1992, *From Morality to Virtue*, Oxford: Oxford University Press; Michael Slote, 1997, "Virtue Ethics", in: Baron/Pettit/Slote, 1997; Michael Slote, 2001, *Morals from Motives*, Oxford: Oxford University Press; and Swanton, 2003.

9 At a time when the claims to distinctness rested on rather vague assertions, and on a seemingly distorted picture of Kantian ethics, it was important to so challenge it, partly to bring out neglected aspects – and the untapped, or undertapped, potential of Kant's ethics. The debate has advanced, thanks in part to rejoinders by Kantians who point out that Kant's ethics does not involve a "big rule" from which all else is deduced; that it is not about rules and actions to the exclusion of virtue and character; that it is not merely about what to do, not about how to be, etc. For one of many works so depicting Kant's ethics, see Julius

More worthwhile, I think, is to examine the roots of the dissatisfaction that has given rise to contemporary virtue ethics. At least, that is what I find more worthwhile: to address the objections put forward by virtue ethicists (or those loosely in their camp) against Kantian ethics, objections that have not gone away. Contemporary virtue ethics was nourished, perhaps even founded on, dissatisfaction with Kantian ethics. Current conceptions of what virtue ethics is do not reflect that dissatisfaction so much as they reflect a wish to offer virtue ethics as a rival to both Kantian ethics and utilitarianism[10]; nonetheless, for some of the leading virtue ethics proponents – in particular, Hursthouse and Slote – the supposed inadequacy of those theories remains a primary reason for developing virtue ethics.

In the remainder of this essay I turn my attention to assessing one very prominent objection to Kantian ethics, the objection that it, along with utilitarianism, places too much value on impartiality. I think it is safe to say that this objection is a major source of the dissatisfaction that gave rise to interest in virtue ethics and is widely held to be a problem besetting Kantian ethics; however, I address it not only for that reason, but also because I find it philosophically intriguing (and because I have already discussed at length elsewhere another objection to Kantian ethics, also philosophically fascinating, that provided some of the momentum for contemporary virtue ethics, the objection that it overemphasizes duty).[11] I'll give particular attention to a distillation of that criticism that is frequently appealed (or at least gestured) to, yet has not been articulated very fully: the "one thought too many" objection.

Moravcsik, 1981, "On What We Aim at and How We Live", in: David Depew (ed.), *The Greeks and the Good Life*, Indianapolis: Hackett University Press, pp. 198–235. For corrections of this picture, see (among others) Marcia Baron, 1995, *Kantian Ethics Almost without Apology*, Ithaca: Cornell University Press; Barbara Herman, 1993, *The Practice of Moral Judgment*, Cambridge: Harvard University Press; Onora O'Neill, 1989, *Constructions of Reason*, Cambridge: Cambridge University Press; and Allen Wood, 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press.

10 Indeed Swanton's particular version of virtue ethics not only is not hostile to Kant's ethics but draws from Kant, as well as from a number of other thinkers. It opposes Kant's ethics primarily in being emphatically pluralistic.

11 In Baron, 1995, and more recently in Marcia Baron, 2002a, "Acting from Duty", in: Allen Wood (ed. and trans.), *Groundwork for the Metaphysics of Morals*, New Haven: Yale University Press, pp. 92–110.

II.

I'll focus initially on the broad claim that impartialist ethics – to use a label often used by critics who favor a (usually neo-Aristotelian) virtue ethics approach – cannot leave room for love and friendship.[12] My hope is to gain a better understanding of the objections (concentrating on them as objections to Kantian ethics) and to determine whether there is a real disagreement between impartialists and their critics, or whether the critics are simply confused or mistaken about Kantian ethics and its implications for love and friendship. I'll also try to defend Kantian ethics against the objections.

Let's begin by looking at the following version of the criticism: Impartialist ethics does not allow us to take the fact that someone is my good friend, or my child, or my sibling as a reason for doing something for that person that I would not generally do for most others. With respect to Kantian ethics, the claim is quite clearly false. (Whether it is false as a criticism of consequentialism is a question that I will not take up here, though I will offer a few remarks on consequentialism in Section IV.) It is perfectly legitimate, on a Kantian view, for me to buy a present for my child that I would not buy for other children, and for me to take the fact that it is my child as a reason for doing so. This is evident from the latitude in Kant's imperfect duties – duties to promote the obligatory ends of others' happiness, and (though of far less relevance to our topic)

12 The claim that Kantian ethics and utilitarianism place too much value on impartiality encompasses more than this objection. Other objections that come under this heading are that the emphasis, or value, placed on impartiality is incompatible with a healthy self-preference or (relatedly) with taking one's own projects seriously or (relatedly) with integrity (though it is very hard to see how this could be a problem for Kantian impartialism) or with resentment (and other moral emotions that have as their objects wrong done to oneself, or misconduct towards oneself, by others). On the objection regarding personal projects and integrity, see the work of Bernard Williams, especially Bernard Williams, 1981a, "Persons, Character, and Morality", in: Bernard Williams, *Moral Luck*, Cambridge: Cambridge University Press, pp. 1–19; and Bernard Williams, 1981b, "Moral Luck", in: Bernard Williams, *Moral Luck*, Cambridge: Cambridge University Press, pp. 20–39; on resentment, see William E. Young, 1998, "Resentment and Impartiality", in: *Southern Journal of Philosophy*, vol. 36, pp. 103–130.

one's own perfection – and in the fact that he recognizes that we have special duties to particular others.[13] Why would anyone think otherwise?

It might be thought that because Kant's ethics enjoins us to respect people as rational beings, we are to like and love them simply as rational beings. Perhaps because everyone is entitled to respect – the stupid along with the brilliant, those who do not utilize their rational faculties very well along with those who do – it is thought that everyone is equally entitled to our affection. Of course that is not Kant's view. We are not enjoined to love or like everyone equally, nor to dole out such resources as material assistance as if we did love everyone equally. The imperfect duty to promote others' happiness does not require (or even encourage) that I promote others' happiness equally, or that I choose which persons to aid impartially, without regard to the fact that some of the people I could aid are close friends or relatives, while others are mere acquaintances, and still others whom I could aid are strangers to me.[14]

But there is more to say about the criticism, and we can explore it more fully while at the same time addressing a closely related objection. The related objection is that mere subjective preference is not given its due in Kantian ethics. That I like someone – even if she is not a relative or a close friend – seems to be a perfectly decent reason for doing something for her that I would not do for just anyone.

---

13 The special duties he recognizes are duties to friends; one surmises that if he had a section in the *Tugendlehre* on familial relationships, as he does on friendship, he would recognize special duties to family, as well. (He does speak of duties of parents towards their children at [*MM* 6: 281] and limited duties of children to their parents, but not of duties of siblings, grandchildren, etc.) It is worth bearing in mind here that, as Allen Wood explains, Kant's *Metaphysik der Sitten* "does not attempt to cover all the ethical duties that we have. This is because Kant confines the 'metaphysics' of morals only to those duties that are generated by applying the principle of morality to human nature in general. But many of our duties, as Kant recognizes, arise from the special circumstances of others, or our relations to them, and especially from the contingent social institutions defining these relations" (Allen Wood, 2008, "Duties to Oneself, Duties of Respect to Others", forthcoming in: Thomas E. Hill Jr. (ed.), *Blackwell Companion to Kant's Ethics*, Oxford: Blackwell).

14 Relevant sections of Kant's works include *MM* 6: 388–394 and 450–454. For discussion, see Marcia Baron/Melissa Seymour, 2008, "Beneficence and Other Duties of Love in the Metaphysics of Morals", forthcoming in: Hill, 2008; Barbara Herman, 2002, "The Scope of Moral Requirement", in: *Philosophy and Public Affairs*, vol. 30, no. 3, pp. 227–256; and Thomas E. Hill Jr., 2002, *Human Welfare and Moral Worth: Kantian Perspectives*, Oxford: Clarendon Press, chs. 4 and 7.

I'll speak to both objections at once – the objection that ties of family and friendship are short-changed, and the objection that mere subjective preference is not given its due – but generally frame my points in terms of subjective preference, since the case might seem easier when it is friendship and family ties that are at stake. There are two questions that we need to ask. First, is the point about subjective preference correct? I.e., is the fact that I like someone a perfectly decent reason for doing something for him or her that I would not do for just anyone? And second, to the extent that it is correct, does it pose a problem for Kantian ethics?

My response to the first question is that with some qualifications, it is correct. The first qualification is that it depends on what it is that I am doing. But when there is a problem – when the fact that I like someone is not a perfectly decent reason for doing something for her that I would not do for just anyone – it is not because liking someone is not an adequate reason for doing something for that person. The problem, rather, is with a failure to do something I should do for others. The problem can emerge in either of the following two forms.

First, the problem might reside in my *not* doing that same thing for *just anyone*. It is permissible for me to invite people to dinner at my home because I like them, and not to invite people over if I do not like them. To offer desperately needed help only to those I like would, however, be far more questionable. The explanation here is pretty straightforward: people, qua people, have a legitimate claim to desperately needed aid, but not to dinner invitations. If the situation is one where aid is desperately needed and not hard or risky for me to provide – e.g., I need only to dial (in the US) 911 – I should aid no matter how little I like the person. The situation is different if the aid is more of a burden to provide and less badly needed (e.g., an acquaintance wants me to baby-sit so that he can go to a movie) because there it is not morally incumbent on me to aid.

Second, the problem might reside in my doing something for someone that precludes my doing something (either the same sort of thing, or something different) for someone to whom I owe a special obligation. The mere fact that I am fond of someone is not a perfectly good reason for bequeathing to him my entire estate if my family has continually provided me with vital support, both emotional and financial, throughout my life. (Let's suppose too that some members of my family are now in considerable financial need, while the friend of whom I am fond is not.) The problem here is not exactly that subjective preference is not by itself much of a reason to favor someone, but rather

that it does not (objectively, and subjectively should not) trump, or even rise to the same level as, reasons for acting in ways incompatible with leaving him my entire estate.[15]

There is an additional qualification (a tacit proviso, in effect) that also should be mentioned. That it is my friend is a legitimate reason for doing *X* for her (even though I would not do *X* for just anyone) – unless it is impermissible to do *X* in the first place. More perspicuously, if an action is immoral, the fact that I am doing it for my friend does not alter its moral status. As William Godwin put it, "What magic is there in the pronoun 'my' to overturn the decisions of everlasting truth?"[16] The fact that someone is my good friend, or my child, or my sibling does not render an immoral action – e.g., aiding him in the commission of a crime – legitimate.

I turn now to the other question, the question about Kant's ethics. Does anything I have said point up a shortcoming in Kant's ethics? Is he unable to accommodate any of the points I have made? Not that I can see. Aiding others falls, in Kantian ethics, under the heading of an imperfect duty, and imperfect duties allow a fair amount of latitude. Although it is hard to say just how much latitude the duty to promote others' happiness allows, we can say this: implicit in Kant's discussion is an expectation that we use good judgment in deciding whom and how to aid, yet there is no prohibition on subjective preference entering in in a situation where no duties to particular others are thereby neglected.[17]

---

15 I thank Justin Brown for this point.

16 William Godwin, 1926, *Enquiry Concerning Political Justice*, ed. by Raymond A. Preston, New York: Knopf, vol. 1, p. 42. Whether Godwin can allow any moral relevance to the fact that this is my sister, etc. is not clear, as I explain in Marcia Baron, 1991, "Impartiality and Friendship", in: *Ethics*, vol. 101, no. 4, pp. 836–857. For discussion of Godwin's infamous example and the morals he draws from it, see, in addition to Baron, 1991, Brian Barry, 1995, *Justice as Impartiality*, Oxford: Oxford University Press; and Susan Mendus, 2002, *Impartiality in Moral and Political Philosophy*, Oxford: Oxford University Press.

17 Perhaps we should add here: provided that in favoring someone simply because one likes him more one does not neglect to aid someone who desperately, urgently needs one's aid, and whom one can easily and safely aid. E.g., one should not put arriving on time to see a dear friend (or anyone else, for that matter) ahead of stopping to phone for emergency assistance for the person lying unconscious on the sidewalk (with no one else on the scene). I say "Perhaps" because it is really hard to know exactly what to add. Kant certainly does not specify any such requirement, but presumably one could not embrace as an end others' happiness and yet regard it as not worth the bother to phone for emergency assistance (or not an adequate reason for being a few minutes late to

Neither subjective preference nor personal ties license one to ignore the needs of others altogether[18] or to ignore duties of gratitude, but this does not render impermissible giving preference in other circumstances to those I like and love.[19]

In sum, Kantian ethics has no quarrel with the view that the fact that someone is my good friend, or my child, or simply a person I like is a legitimate reason for doing something for that person that I would not do for just anyone, provided that it is understood that neither subjective preference nor personal ties license me to ignore the needs of others altogether, neglect urgent, grave needs, or ignore other duties, such as duties of gratitude.

III.

Partiality – whether due to mere subjective preference or to ties of family or friendship – is permissible within moral bounds. It is fine to act for the reason that this person is my friend, my brother, even my favorite student; but the fact that this person is my friend, etc., does not render an immoral action permissible, and that I am helping my friend does not alter the fact that the needs of others also make a moral claim on me.

Favorite student? Can an impartialist regard as permissible having a favorite student? I don't see why not. Sometimes it is supposed that partiality is permissible except with respect to those to whom one stands in a certain sort of relationship, e.g., one's students. But things don't

meet a friend). My uncertainty about what to put in the proviso arises from the fact that at almost any given time we could help some very needy person somewhere, and do so easily and safely (by, say, contributing money to Oxfam); yet it seems clear that Kant's duty of beneficence does not require that we do so at every such opportunity. For a searching discussion of Kant on beneficence, see Herman, 2002.

18 This seems to be the point of *MM* 6: 390. As I explain in Baron, 1995, Ch. 3, Kant's point seems to be that we may not altogether reject an end of (for example) helping needy strangers on the ground that we are taking care of elderly parents; but we may limit our pursuit of the first end in favor of the second, putting much more energy into the latter and relatively little into the former. However, there is room for debate about just how to read this passage. See Baron/Seymour, 2008.

19 If Kant's ethics is to be criticized here, it would make more sense to criticize it for allowing subjective preference so much play, than for allowing it too little. There is certainly scope for objections to the effect that we should in fact be required to attend especially to greatest need.

divide up in this way. Depending on the behavior in question, it is permissible to behave towards my students in a way that reflects differential "liking" for them, inviting only those I like to dinner at my home, for example; yet it is not okay for the grades I give to express differential liking, nor would it be permissible for me to agree to write letters of recommendation only for those I like, rather than for all those I think are good enough as students to merit a recommendation. (Of course some professional responsibilities are such that one has to be very careful indeed not to show partiality. Psychiatrists and clinical psychologists need to take care not to show partiality towards certain patients, but then they are also barred from forming friendships with their clients, something not barred to professors and their students.) Nor, more importantly, is it the case that outside of professional contexts, partiality is never a moral issue.[20] My fondness for my friend will properly express itself in all sorts of things that I do for her but do not do for others, but it must recognize moral limits. The limits are simply the moral limits on all conduct. If it is immoral to cheat on a test, and to help someone to cheat on a test, the fact that I am helping my friend cheat and helping him out of affection for him does not justify my action. It shows me to have some admirable character traits (along with some not so admirable ones), but neither my affection for my friend nor our friendship renders permissible my act of helping him to cheat.

This might conceivably be a point on which partialists take issue with impartialism. I doubt it, but I might be wrong. Susan Wolf remarks that the thought that there is nothing wrong with someone who violates impartial morality to protect her son or daughter is "perhaps the strongest motivating thought behind partialist morality". Her example is of a woman whose son has committed a crime and who chooses to hide him from the police, even though she knows that unless he is caught, an innocent person will be wrongly convicted and imprisoned for the

20 I emphasize this because it is often supposed that impartiality is needed only in special contexts, e.g., when one is grading papers, sitting on a jury, or otherwise in the seat of a "judge," or more generally, that it is needed in what might be thought of as "public" contexts but not in "private" ones. Even apart from the problems with a public/private distinction, this view is misguided. We need to be sensitive to considerations that call for a perspective of impartiality rather than assume they are morally irrelevant except in certain spheres of activity and that apart from that, there is no need to take up an impartial perspective (or worse, that it is usually inappropriate to do so). See Baron, 1991.

crime.[21] It seems to me, however, that very few partialists would think it morally justifiable for the mother to hide him from the police, but they might claim that though the action was wrong, there is much in it, or behind it, that is admirable. If this is their view, they are not in disagreement with Kantians (and other impartialists). The Kantian could make the same claim – though Kantians are probably slower to find something admirable in an immoral action committed out of love than are their partialist critics.

It is more likely that any partialists who find the position objectionable – the position that the fact that the person is my child or my very good friend does not render an immoral action morally permissible – do so because they misunderstand it. It is easily confused with the position that the fact that someone is my good friend or my child could never be relevant to the question of whether something is morally permissible. But they are different, and the first position does not entail the second. It might well be the case that paternalism of certain sorts is permissible for me with respect to my child but not with respect to anyone else (including other children). It might well be the case that pouring out my woes is permissible vis-à-vis certain people who are very close to me, but would be an imposition on others. And it is plausible to claim that it would be wrong of me never to throw a birthday party for my child, but not wrong of me never to throw a birthday party for any other child.[22] Almost all impartialists would agree with these claims, and thus would agree that the fact that it is my friend or child can be relevant to whether something is permissible.[23]

---

21 Susan Wolf, 1992, "Morality and Partiality", in: *Philosophical Perspectives*, 6, Ethics, pp. 243–259, at pp. 253–254. She suggests that it might well be reasonable for the woman to hide him and that some might infer from this – though she does not – that it must be moral, too. Wolf's view is that it is not moral, and that the conflict here is between morality and the demands of love rather than between competing moral concerns.

22 This of course depends on the circumstances of my life; if I live with a friend and her child, and the friend has to be away on business at the time of the child's birthday, refusing to throw a birthday party for the child because it isn't my child would be highly dubious. But the point is clear enough: the wrongness of not throwing a party has to do with the nature of the relationship between the child and me. I thank Wade Robeson for pressing me on this point.

23 Critics of impartialism sometimes write as if it is part of impartialism to disagree with these claims (and to hold that the fact that it is my friend or child can have no bearing on the permissibility of my action). The only authors I am aware of who come close to fitting their picture of impartialists are William Godwin and

Of course it will often take sensitivity and finely-tuned judgment to determine whether the special relation one bears to the person is of a sort to warrant acting as one proposes. It is not as if the mere fact that this is my child justifies a particular paternalistic action that I am thinking of taking; the particulars about the child, and not the mere fact that he is my child, matter, as do the particulars of our relationship. Knowing to which friends it is fine (and perhaps even important to the relationship) to pour out my innermost doubts and fears likewise requires sensitivity and perceptiveness.[24]

This need for sensitivity and finely-tuned judgment might seem to pose a problem for Kantians. Some critics hold that it is part of Kantian impartialism that principles, and the commitment to adhering to them, are supposed to supply all of the motivation and understanding that the agent needs. Arguing that compassion and other "care virtues cannot be captured within the framework of … impartialist morality," Larry Blum claims that according to impartialism, there is no need for agents "to draw on moral capacities beyond the mere ability to consult the principle."[25] This may be true of the impartialism of Lawrence Kohlberg (the focus of Blum's discussion), but it is not characteristic of Kantian impartialism in moral philosophy. And Kant himself held that the moral law requires "judgment sharpened by experience" for its application [*G* 4: 389] and shows in many passages in the *Tugendlehre* that sensitivity is crucial for moral conduct. For example, in discussing beneficence, he writes:

> Someone who is *rich* (has abundant means for the happiness of others, i.e., means in excess of his own needs) should hardly even regard beneficence as a meritorious duty on his part, even though he also puts others under obligation by it… He must also carefully avoid any appearance of intending to bind the other by it; for if he showed that he wanted to put the other under an obligation (which always humbles the other in his own eyes), it would not be a true benefit that he rendered him [*MM* 6: 453].

Peter Singer. (I think it no accident that the only people who even come close to taking this view are consequentialists.) See notes 16 above and 36 below.

24 The idea here is (a) if an action really is immoral, the fact that it is being done for a loved one does not make it permissible; but at the same time (b) there are actions whose permissibility or impermissibility depends on the relationship the agent bears to the recipient. These usually are actions done for the benefit of the recipient or for the relationship itself (or perhaps for the agent) *and* are actions that (unlike cheating, or violating someone's confidence) have no significant negative bearing on others.

25 Lawrence Blum, 1994, *Moral Perception and Particularity*, Cambridge: Cambridge University Press, p. 203.

That it is a duty (albeit indirect) to "cultivate the compassionate natural (aesthetic) feelings in us, and to make use of them as so many means to sympathy based on moral principles and the feeling appropriate to them" [*MM* 6: 457] is further evidence that sensitivity is crucial, in Kant's view, for moral conduct. It should also be borne in mind that there are, according to Kant, "certain moral endowments" that "lie at the basis of morality, as *subjective* conditions of receptiveness to the concept of duty," viz. "*moral feeling, conscience, love* of one's neighbor, and *respect* for oneself (self-*esteem*)" ([*MM* 6: 399], italics Kant's).[26]

In sum, the fact that sensitivity and judgment will often be needed to determine whether the special relation one bears to the person warrants acting as one is thinking of acting is not a problem for Kantian impartialism, for Kantian ethics does not hold that such capacities are unnecessary.

IV.

It seems so far that there is very little in Kantian impartialism for partialists to disagree with. So what *is* the basis for their charge of excessive impartialism? One possibility is that the main, or most serious, objections that partialists have and that they speak of as applying to impartialist ethics in general actually apply only to consequentialism. I bring this up because it seems to me that the strongest argument that claims that impartialist ethics does not leave room for love and (deep) friendship is one that does not apply to Kantian ethics and has greater plausibility as a criticism of consequentialism.

Here is how the criticism goes if it is a criticism of consequentialism. Consequentialism requires us to maximize some impersonal value. (I am here ignoring satisficing consequentialism, and will throughout my paper.) But this requires putting impersonal value ahead of our friends and loved ones, even if it is just a slight increase of impersonal value that can be realized by, say, taking on a major project aimed at alleviating homelessness or world hunger, a project that requires that we spend far

26 For further discussion, see Baron, 1995, Ch. 6; Lara Denis, 2000, "Kant's Cold Sage and the Sublimity of Apathy", in: *Kantian Review*, vol. 4, pp. 48–73; Paul Guyer, 1993, *Kant and the Experience of Freedom*, Cambridge: Cambridge University Press, ch. 10; Nancy Sherman, 1997, *Making a Necessity of Virtue: Aristotle and Kant on Virtue*, Cambridge: Cambridge University Press, ch. 4; and Wood, 1999.

less time with our young children than we otherwise would. Consequentialists retort that consequentialism does not require that we base our decisions on a calculation of what will maximize impersonal good. Making decisions on that basis is likely to reduce, on the whole, how much impersonal good is produced. More good is produced if in fact we do not focus on maximizing impersonal value and instead rely on common-sense moral "rules." They may even claim that it is best that we *never* assess our options by asking what will maximize utility, or some other impersonal value, and indeed that it is best if we never hear of consequentialism at all.

This may seem to be a viable solution, but let's take a closer look. The consequentialist has two options. The first is to say that consequentialism is intended only as an analysis of what rightness consists in and perhaps as a tool for determining retrospectively whether what we or someone else did was right (though once it is introduced as a useful tool, it seems likely that people would want to use it prospectively as well; so probably introducing it as a tool of any sort would be too risky). But if consequentialism has no implications for how to lead our lives, it is not a very helpful theory. At best it would be useful only for an elite, who would have the authority to decide what "moral rules" to try to manipulate (almost) everyone into accepting as "commonsense" (and to decide which moral rules that they now accept they ought to be persuaded to discard). This would of course work only if social and political arrangements were such as to invest in some people the power to so manipulate the populace. It would not work in a society in which the free exchange of ideas was encouraged. For it to work, moral philosophy would be "classified" research, not a subject to be offered to students; or at the very least, consequentialism would itself have to be entirely off the table, rather as communism is in the United States. Public policy debates would have to be fully concealed from the public.

The other option – an option that is necessary if consequentialists want their theory to provide guidance on how to live, and not through having a secret elite, invested with power to manipulate our moral views – is to acknowledge that there have to be some points at which we assess what we are doing and make adjustments so as to maximize impersonal good. But then we are back (almost) where we started, for doing this involves putting impersonal good ahead of loved ones, even if it is just a slight increase of impersonal value that can be realized by neglecting loved ones; and it is in this sense that consequentialism is incompatible with love and friendship.

I do not mean to endorse this as a knock-down argument, but I think it is serious as a criticism of consequentialism. (Peter Railton goes quite far in addressing the objection, however, in his "Alienation, Consequentialism, and the Demands of Morality").[27] By contrast, as a criticism of Kantian ethics it makes no sense at all, since Kantian ethics has no requirement that we maximize impersonal value (or, for that matter, any value). Kantian ethics is far less concerned with bringing about certain results than is consequentialism, and to the extent that it is concerned about it – we are, after all, obligated to promote others' happiness – there is no requirement that we do as much as we possibly can to bring about such results. We have a duty to promote others' happiness (and to perfect ourselves), but just how we go about this is to a considerable extent up to us. There is no requirement that we maximize others' happiness (or the general happiness, or respect for persons, or any other good). A Kantian's positive duties thus appear less likely to dominate her life in a way that precludes or undermines meaningful relationships than do a consequentialist's positive duties.

One possibility, then, is that some critics fail to notice that their argument against impartialism relies on the assumption that impartialism is concerned to maximize impersonal good, an assumption that is true only for consequentialist, and not for Kantian, impartialism. Actually, it is more than just a possibility. In his reply to Railton's paper, William Wilcox says that his criticism of Railton applies to my work, as well, apparently not noticing that his criticism hinges on the theory under criticism being of a sort that requires maximizing impersonal good.[28]

Another possibility is that partialist critics may maintain that while Kantians *want* to hold that it is permissible to do favors for a loved one that one would not do for most others – and permissible to do so because of who she is, and because of one's love for her – in fact they cannot consistently take this view. I don't know of anyone who has *argued* this, though some have assumed that there is an inconsistency. The belief that

27 Peter Railton, 1984, "Alienation, Consequentialism, and the Demands of Morality", in: *Philosophy and Public Affairs*, vol. 13, no. 2, pp. 134–171. For critical discussion, see William Wilcox, 1987, "Egoists, Consequentialists, and their Friends", in: *Philosophy and Public Affairs*, vol. 16, no. 1, pp. 73–84; Dean Cocking/Justin Oakley, 1995, "Indirect Consequentialism, Friendship, and the Problem of Alienation", in: *Ethics*, vol. 106, no. 1, pp. 86–111; and (more supportively) Alastair Norcross, 1997, "Consequentialism and Commitment", in: *Pacific Philosophical Quarterly*, vol.78, no. 4, pp. 380–403.

28 Wilcox, 1987, pp. 73–74.

Kantians cannot consistently take this view is fueled, I suspect, by a confusion concerning the "level" at which an impartial standpoint is required, on an impartialist picture.[29] Whereas impartialists hold that justification has to take place from an impartial standpoint, critics often suppose that impartialists believe that one should *always* take an impartial standpoint. So, whereas impartialists hold that doing *X* (say, lavishing one's resources on one's children) is justifiable only if it would be permissible for anyone similarly situated to do *X*, it is often claimed that impartialists hold that the mere fact that it involves favoring certain people over others, and favoring them simply on grounds of personal attachment or family ties, suffices to render an action unjustifiable.[30] What is involved here is a level confusion. At a meta-level – at the level at which we reason about what is permitted – we are, on an impartialist

29 In Baron, 1991. See also Cynthia Stark, 1997, "Decision Procedures, Standards of Rightness, and Impartiality", in: *Nous*, vol. 3, pp. 478–495. For misgivings about a two-level approach, see Susan Mendus, 2002, *Impartiality in Moral and Political Philosophy*, Oxford: Oxford University Press, pp. 59–64.

30 John Cottingham argues: "When I sit up all night with my sick child, the impartiality thesis tells me that I am not acting morally; or at least, if my action is to be justified morally, I have to show that I could not be making a greater contribution to human welfare by helping any other child who may be in greater need of care and attention" (John Cottingham, 1983, "Ethics and Impartiality", in: *Philosophical Studies*, vol. 43, pp. 83–99, at p. 88). The impartiality thesis holds that "to reason ethically, to consider things from the moral point of view, is to adopt an *impartial* standpoint." It is reasonable to suppose that impartialists hold this thesis (although I should note that I, at least, would count as ethical reasoning reasoning about what is, say, best for my child, as long as in thinking about this I recognize and am prepared to abide by moral constraints on what it is permissible to do for my child). But does it entail what Cottingham says it entails? Only if it is interpreted in a way that virtually no impartialist would accept. The unclarity concerns at what level, or with regard to what sorts of things, one is to adopt the impartial standpoint. Is it with regard to actions one is thinking of performing? Or is it with regard to the principles which hold some promise of giving one direction as to what actions to take? In the first instance, the idea would be that to reason ethically about whatever it is that I am thinking of doing, I must take care not to allow my decision about what to do to be influenced by partiality to myself or to others. In the second instance, the idea would be that we should, in deciding whether it is permissible for us to do *X*, detach ourselves as far as possible from our own particular attachments and ask whether it is permissible for anyone similarly situated to do *X*. Impartialists hold the latter, not (usually) the former; but the principle Cottingham adduces has the implication he alleges only if it is understood to apply as described in the first instance (as well as in the second instance). For more on this, see Baron, 1991 and Stark, 1997.

view, to be strictly impartial: the principle that I endorse for others, I endorse for me and my kin, too. If something is permissible for me, it is permissible for everyone similarly situated. If it is wrong for children in circumstances *C* to do *X*, then it is wrong for my kid, who is in circumstances *C*, to do *X*. So, impartialists do endorse the irrelevance of "But it's my kid we're talking about!" to the validity of the principle in question (and to its reach to my child). But that is *not* to endorse the idea that "But it's my kid!" is never morally relevant. Some critics have wrongly supposed that it is.[31] They have not recognized that the position that in deciding whether a certain principle or policy is permissible, I should ask whether it is permissible for anyone similarly situated, does *not* entail that I should never allow my decision about how to act to be influenced by partiality to, or my bonds to, certain others. Now *if* one thought that my decisions about how to act should always be dictated by moral principles, there would be little need to distinguish impartiality at the level of principles from impartiality at the level of deciding what to do. (There would still be a need to draw such a distinction, though; it would be between the content of the principle, which could include particular ties, as in "Honor your parents," and reasoning about, and seeking to justify or to challenge, such a principle). But if, as is the case in Kantian ethics, moral principles typically limit what is permissible without directing us to take precisely this option and no other, there is ample room for subjective preference and ties of love and friendship to influence one's choices without coming into conflict with morality.

So far I have canvassed some reasons for objecting to "impartialism" or "impartialist ethics" and have said that at least as applied to Kantian ethics, the objections are based on a confusion or a misunderstanding. But I do think there is probably a genuine disagreement to be ferreted out, so I'll probe further.

V.

Susan Wolf's discussion of impartialism in her essay, "Morality and Partiality," provides another point of entry into determining just what the partialist objections are to impartialism (and in particular to Kantian impartialism). It is admittedly an odd point of entry, for her characterization of impartialism is quite sweeping. Her characterization centers on

31 See for example Cottingham, 1983.

what she calls the "Impartialist Insight." The Impartialist Insight is "the claim that all persons are equally deserving of well-being and respect." Impartialism is "the position that a moral person is one who recognizes and appreciates the Impartialist Insight and integrates it into her life."[32]

Although Wolf did not intend this as a slap in the face to partialists – indeed, she expresses a strong sympathy for their views – the characterization might well offend. Partialists might protest that they too recognize that all persons are equally deserving of well-being and respect, and that they too hold that a moral person recognizes and appreciates this insight. Depending on what is meant by "integrating the insight into one's life," aren't virtually all of us impartialists? I'll set aside this concern and consider how her very inclusive characterization can help locate disagreement about the value of impartiality and its place in our lives. Lest this seem offensive or, if not that, simply bizarre, I should explain that as I see it, the division between impartialists and partialists, and the labels "impartialist" and "partialist," are clumsy, and thus it is not imperative that a characterization or definition of impartialism accurately capture that division. Since my concern is to locate, understand and attempt to reply to the criticisms of Kantian impartialism – criticisms that are at the heart of the motivation for virtue ethics – rather than to find a characterization of impartialism that does justice to the labels "impartialists" and "partialists," Wolf's characterization is useful for my purposes. And while it has the result that some who would usually be called "partialists" become, on her terminology, "moderate impartialists," it does provide for a wide range of views regarding the proper role or place of impartiality in our lives.

I take it that her characterization of impartialism is not meant to entail that according to impartialists, it is a sufficient condition of being a moral person that one recognize and appreciate the Impartialist Insight and integrate it into one's life. The idea, presumably, is that it is a necessary condition. (Any doubts one might have about this are allayed by her remarks to the effect that morality is "not just about treating people equally or fairly, but about treating them well.")[33] That being the case, room for disagreement – and for varieties of impartialism – will arise in connection with what it means to integrate the Impartialist Insight into one's life (and relatedly, just what it is to appreciate the

32 Wolf, 1992, p. 245.
33 Ibid., p. 247.

Insight). As Wolf says, integrating it into one's life need not involve letting it absolutely take over.[34]

What will it involve? At a minimum, that one holds herself to "the same standards that she expects of others." This is the "strict" part of impartialism. The part that is harder to spell out precisely is this: integrating the Impartialist Insight into one's life means that one is "moved to practical effect by the thought that others – all others – are as deserving of the fundamental conditions of well-being and respect as are she and her circle of friends and loved ones." It will have some practical effect, as Wolf puts it, "on one's politics, on one's activities, on one's choice of how to spend one's money."[35]

The place where disagreement initially seems most likely is in the last bit I quoted – in how much and what sort of effect one's appreciation of the Impartialist Insight will have on one's politics, one's activities, one's choice of how to spend one's money. But I am not sure that such disagreements will have much to do with disagreements about the value of impartiality and the value of bonds of love and friendship. They will have more to do with views about property rights, economic distribution, etc., views which could, but needn't, be tied to views about the value of impartiality and the value of family bonds. I am not aware that impartialists more typically hold that one should give a large percentage of one's income to charity or to social programs designed to reduce social and economic inequalities than do partialists. It is true that impartialists may have a harder time justifying gross economic disparities, but there are many different views as to the best ways – and the prospects for – altering the disparities, and it does not seem that partialists and impartialists line up on opposing sides. So far then, it does not look as if disagreements regarding the practical effect that an appreciation of the Impartialist Insight will have on one's life will be such as to split partialists and impartialists into two camps, or even into several groups that could be plotted along a scale where extreme impartialism is at one end, and extreme partialism is at the other.

But there is another way in which Wolf's characterization of impartialism yields a source of disagreement. The disagreement concerns how one is to integrate the Impartialist Insight into one's life but does not mainly concern the practical effect of the Insight on one's politics and on how one chooses to spend one's money. It has more to do with how it is

34 Ibid., p. 245.
35 Ibid., p. 246.

that we are to integrate the Impartialist Insight into our lives without letting it "absolutely take over." Specifically, it concerns *just how ready we should be to take up a detached perspective, to view our relations to others just as someone's relations to others (and to think of the people we love just as some persons*). It concerns the notion that there can be "one thought too many."

I can best reflect the sorts of disagreements I have in mind by noting some disagreements among those who write on the subject. First, consider Wolf's rejection of some forms of impartialism. In characterizing impartialism as she does, Wolf makes a point of avoiding what she calls "Extreme Impartialism." One form of Extreme Impartialism is something that virtually all impartialists would reject: on this view, "a person is morally required to take each person's well-being, or alternatively each person's rights, as seriously as every other, to work equally hard to secure them, or to care equally much about them, or to grant them equal value in her practical deliberations."[36] That the view is absurd is evident from the fact that it entails, as she puts it, that "morality requires one to care, or to act as if one cares, no more about one's own child than about a stranger's." So far I am in complete agreement with Wolf. But she deems "only slightly less absurd" the "much more popular view … that permits partisan emotions and behavior, as long as in fact they promote nonpartisan goals."[37] Apart from the fact that her wording makes the connection between "partisan emotions and behavior" and "nonpartisan goals" sound rather loose, this view is the standard consequentialist reply to partialist criticisms. It is, I take it, Railton's position. (In fact as stated it is weaker than the standard consequentialist reply because it says nothing about maximizing: it does not require that the partisan emotions and behavior *maximally* promote nonpartisan goals. But her subsequent comment on it does mention maximizing).

Why does Wolf think it extreme, and rather absurd? She explains: "The acceptability of coaching one's daughter's soccer team, or taking one's friend to dinner on her birthday does not rest on the fortuitous coincidence that this action, or even the way of life that gives rise to it, is

36 Ibid., p. 244. One impartialist who appears to accept this thesis is Peter Singer. He tries to argue from the position that ethical judgments should be made from a universal point of view to the conclusion that ethical conduct requires agents to weigh up the interests of "all those affected" by one's decision and "adopt the course of action most likely to maximize the interests of those affected" (Peter Singer, 1979, *Practical Ethics*, Cambridge: Cambridge University Press, pp. 10–12).

37 Ibid., p. 244.

the one that will maximize human welfare or equal respect all around." Not, I agree, on the *fortuitous coincidence*, but if we replace "fortuitous coincidence" with "fact" is the view that she rejects absurd? Not being a consequentialist, I do not endorse the view; but it does not seem to me to be an "extreme" form of impartialism, or in any way absurd.[38] I suspect that to those who find it absurd, the problem is that it runs afoul of the "one thought too many" requirement, a requirement that I find elusive, but which seems to involve the notion that justification either should not be sought, or should not be sought from a distant, objective point of view. More on that shortly.

VI.

The first disagreement was Wolf's disagreement with a view that is associated with Railton. The second is a disagreement I have with Railton. Whereas Wolf finds extreme impartialism in the view he develops, I find no such thing; by contrast, I find him to be too hard on impartialism (and in that sense, insufficiently impartialist). Railton's article begins with two examples that are designed to show the need to explain how impartial morality and friendship are compatible. For the sake of brevity, I'll discuss just one of them, but since the details matter, I need to quote it in full.

> To many, John has always seemed a model husband. He almost invariably shows great sensitivity to his wife's needs, and he willingly goes out of his

38 An alternative is suggested by Cottingham (John Cottingham, 1997–1998, "The Ethical Credentials of Partiality", in: *Proceedings of the Aristotelian Society*, vol. XCVIII, pp. 1–21). Cottingham emphasizes that "goodness grows *from the inside outwards*," i.e., that "human lives are valuable not primarily in virtue of how far they conform to impersonally defined rules of conduct, or in so far as they contribute to some giant amalgam called 'the good', but in so far as they are lived in ways which give richness and meaning to the short journey each of us has to undergo" (pp. 7–8). And thus one might argue that it is absurd to think that the acceptability of coaching one's daughter's soccer team and the like rests on the fact that this action or the way of life that gives rise to it will maximize human welfare for, the argument might go, they rest on the meaning they give to the agent and those close to her. But this is implausible; for even if we grant Cottingham's point about the source of value, the question of permissibility or acceptability still remains, and it hardly seems absurd to think that the permissibility rests on facts about how this action fits or fails to fit with something beyond it, e.g., whether it or the way of life that gives rise to it maximizes human welfare.

> way to meet them. He plainly feels great affection for her. When a friend remarks upon the extraordinary quality of John's concern for his wife, John responds without any self-indulgence or self-congratulation. "I've always thought that people should help each other when they're in a specially good position to do so. I know Anne better than anyone else does, so I know better what she wants and needs. Besides, I have such affection for her that it's no great burden – instead, I get a lot of satisfaction out of it. Just think how awful marriage would be, or life itself, if people didn't take special care of the ones they love." His friend accuses John of being unduly modest, but John's manner convinces him that he is telling the truth: this is really how he feels.[39]

John's reply is supposed to be jarring, unsettling. It may be unwise of me to admit this in print, but I don't find it jarring. Not that it strikes me as entirely devoid of oddness. The one thing that I find odd in his reply is something that was pointed out by one of my students, who remarked that she kept waiting to learn that Anne was a paraplegic. John's tone regarding his wife seems peculiarly solicitous. But I don't think that has anything to do with what Railton expects us to find disturbing.

What might be thought a more appropriate reply to his friend? John could simply say that he loves Anne, but this would not speak to his friend's suggestion that there is something remarkable about the quality of John's concern. Presumably lots of husbands love their wives. And it would be presumptuous of John to suggest that he loves his wife more than most husbands love their wives, or that Anne is exceptionally lovable, and that that is why he does more than most husbands do for their wives. Alternatively, John could claim that he does not think that he is unusual, but that if he is, that suggests that something is amiss with other husbands. That reply, I suppose, would be just as "impersonal" as the one that he does offer.

I don't see John's reply to be disturbingly impersonal, and, apart from the solicitude that I noted, I don't see it as in any way peculiar. Why does Railton think that it is? He invites us to reflect on how John's remarks might sound to Anne.

> Anne might have hoped that it was, in some ultimate sense, in part for her sake and the sake of their love as such that John pays such special attention to her… It is as if John viewed her, their relationship, and even his own affection for her from a distant, objective point of view – a moral point of view where reasons must be reasons for any rational agent and so must have an impersonal character even when they deal with personal matters. His wife might think a more personal point of view would also be appropriate, a

39 Railton, 1984, p. 135.

> point of view from which "It's my wife" or "It's Anne" would have direct and special relevance, and play an unmediated role in his answer to the question "*Why* do you attend to her so?"[40]

An answer to that question, it seems to me, would be likely to take the form of "I think relationships are best if…" Actually, such questions are hard to answer, partly because it is not clear just what the person is wondering. But a reasonable guess would be that John's friend is not wondering what particular qualities of Anne's make her especially lovable, or anything of that sort, but rather is interested in hearing the sort of thing that John in fact does say. Later in his essay, Railton draws an analogous example, of Juan and Linda, and what is interesting is that the chief difference between Juan and John is that it takes more time, and some very definite prodding from the questioner, to get Juan to take a distant, impersonal perspective. This supports my hunch that one difference in views about impartialism concerns *how ready* one can be to take up the distant, impersonal perspective without that readiness properly raising suspicions. Juan is less ready than John; that, it seems, is why he, unlike John, does not seem to Railton to be alienated from his wife or from his affections. So perhaps one disagreement between partialists and impartialists, and between moderate impartialists and immoderate or extreme impartialists concerns how readily one should be able to shift from a personal to an impersonal perspective.

There is a complication that I omitted. Discussing John, Railton says, "That he devotes himself to her because of the characteristically good consequences of doing so seems to leave her, and their relationship as such, too far out of the picture…"[41] Now, if John does devote himself to her *because* of these consequences, then I agree with Railton (though I didn't read John's reply to his friend in this way). But Railton also says, as quoted earlier, that the answer that Anne might hope for, and which he thinks is in order, is one in which "It's my wife" or "It's Anne" would play *an unmediated role.* And the suggestion thus seems to be that the thought that marriages are better, and indeed that the world is a better place, thanks to such devotion, is one thought too many.[42]

---

40 Ibid., p. 136.

41 Ibid., p. 136.

42 Related to this position is the claim that actions done from love or friendship should not be "mediated" (or rather, the love itself should not be mediated) by a commitment to doing what is right, being a virtuous person, or anything else of that ilk. Here again the idea seems to be that one thought too many is involved –

Now, to connect this up with the question of why impartialism is found objectionable: It seems to me that the deepest disagreement that partialist critics have with impartialism (once misunderstandings are cleared away), and that some moderate impartialists have with what they might, following Wolf, regard as "extreme impartialism," concerns when, and how readily, it is decent or even good to take up a detached perspective, and when doing so involves having "one thought too many." That being the case, it is now time to try to figure out what it means to have one thought too many, in the way developed by Bernard Williams and endorsed by (among others) Julia Annas and Susan Wolf and to some extent by Peter Railton.[43] (Given my leanings towards rather extreme impartialism, I am not in the best position to appreciate the possibility of one thought too many. Hopefully this essay will prompt some who are in a better position to explain it.)

VII.

I do understand that there are thoughts such that the fact that one has them calls the person's character into question – thoughts, for example, about what forms of torture one would most enjoy inflicting, should one

---

even, it seems, if the person is not thinking about rightness or virtue while acting lovingly. I discuss this below, and at length in Baron, 1995, ch. 4.

43 The locus classicus of "one thought too many" is Williams, 1981a. In mentioning Julia Annas, I have in mind Julia Annas, 1984, "Personal Love and Kantian Ethics in Effi Briest", in *Philosophy and Literature*, vol. 8, pp. 15–31. I reply to her article in Marcia Baron, 1988, "Was Effi Briest a Victim of Kantian Morality?", in: *Philosophy and Literature*, vol. 12, pp. 95–113. Wolf invokes the notion of "one thought too many" in Susan Wolf, 1982, "Moral Saints", in: *Journal of Philosophy*, vol. 79, no. 8, pp. 419–439, as does Jean Rumsey in Jean Rumsey, 1997, "Re-visions of Agency in Kant's Moral Theory", in: Robin May Schott (ed.), *Feminist Interpretations of Immanuel Kant*, University Park: Pennsylvania State University Press; Harry Frankfurt also endorses it, though he takes issue with Williams' way of putting the point (Harry Frankfurt, 2004, *The Reasons of Love*, Princeton: Princeton University Press, pp. 36–37). For (generally critical) discussions, see Baron, 1995, ch. 4; Guyer, 1993, ch. 10; Herman, 1993, ch. 2; Thomas Nagel, 1982, "Williams: One Thought Too Many", in: *The Times Literary Supplement*, May 6, reprinted in: Thomas Nagel, 1995, *Other Minds*, Oxford: Oxford University Press, pp. 167–173; Madison Powers, 1993, "Contractualist Impartiality and Personal Commitments", in: *American Philosophical Quarterly*, vol. 30, pp. 63–71; Samuel Scheffler, 1992, *Human Morality*, Oxford: Oxford University Press; and David Velleman, 1999, "Love as a Moral Emotion", in: *Ethics*, vol. 109, pp. 338–374.

find oneself in a position where torturing others was part of one's job description.[44] Likewise, if upon reading Bernard Williams' example of Pedro and Jim, one's thought was "I'd never ordinarily choose to kill an innocent person, but if I were in Jim's situation, not only would I not hesitate, I would savor the richness of the experience."[45] But even if, as seems unlikely, one or both of these thoughts should be classified as *one thought too many* – as opposed to something like *one horrendous thought* – they do not seem to be a helpful model for understanding how it is that the thought that what one is doing is morally right is one thought too many.

I also appreciate that there are times when some thoughts, though innocuous enough in themselves, are inappropriate, and cause for some consternation. I remember a film, "Ordinary People," in which the protagonist recalls that as he and his wife dressed to go to the funeral of their child, she suggested that he wear a different pair of shoes (and not for reasons of comfort). That she was thinking at that time about which shoes he should wear troubled him, and possibly it is in roughly this way that some find attention to moral matters at a time such as when one's spouse is in grave danger of drowning, disturbing. But the point about one-thought-too-many seems to go beyond this. And it seems to take two rather different directions. The more radical direction is that justification is unnecessary or that justification should not be sought from an impersonal, impartial perspective; this I take to be Williams's suggestion. (In reference to the example of someone in a position to save one but not both people in an accident, and who chooses to save his wife, Williams says that the "random" element in such events…should be seen…as…a reminder that some situations lie beyond justifications.")[46] The other, more moderate direction is the one that I see in Railton's discussion: Railton is not at all opposed to seeking such justification but has doubts about agents who are quick to think about their deepest attachments from an impartial perspective. A readiness to detach and view things objectively is, it seems, alienating, or expressive of alienation, alienation

44 I discuss this in Marcia Baron, 2002b, "Character, Immorality, and Punishment", in: Walter Sinnott Armstrong/Robert Audi (eds.*), Rationality, Rules, and Ideals: Essays on Bernard Gert's Moral Theory*, Lanham: Rowman and Littlefield, pp. 243–258.

45 See Bernard Williams, 1973, "A Critique of Utilitarianism", in: J. J. Smart/ Bernard Williams, *Utilitarianism: For and Against*, Cambridge: Cambridge University.

46 Williams, 1981a, p. 18.

both from others and from one's own affective nature. At least this seems to be the idea. The more moderate direction, taken by Railton, may just be that there are times when such detachment is inappropriate; but I expect, given his examples, that there is more to it than this.

For other (arguably more extreme) misgivings about one-thought-too-many, I turn to a discussion by Julia Annas, "Personal Love and Kantian Ethics in *Effi Briest*," in which she argues that whether or not it was his intention, Theodor Fontane's novel, *Effi Briest,* shows the disastrous effects of a Kantian moral outlook.[47] The relevant part of the story is as follows. When Effi's husband learns that six years earlier she had an adulterous affair, he divorces her (and, challenging the former lover to a duel, kills him), and Effi's parents announce that they will have nothing to do with her. Subsequently their family physician, concerned about Effi's failing health and general misery, urges her parents to drop their stance towards her and invite her to live with them. Her mother is initially reluctant and says to her husband, "I love her as much as you do, perhaps more. But we've not been sent into the world just to be weak and forebearing and show respect for all that's against the laws of God and man." He replies: "Oh really, Luise. One thing's more important…Parents' love for their children."[48]

What Effi's father says seems to me straightforward: he is saying that parents' love for their children (and expressing that love) are more important than showing respect for prohibitions on adultery by refusing to have anything to do with anyone who ever committed adultery. He is speaking from an impartial point of view – and, I would add, doing so quite appropriately. Annas rejects this reading (or something very like it).[49] She writes: "He does not mean that parents are morally permitted to love their children, which could hardly have much weight against their moral condemnation of Effi's actions, and anyway would be a classic case of what Williams aptly calls having 'one thought too many,' weakening

47 See Annas, 1984.

48 Theodor Fontane, 1967, *Effi Briest*, trans. by Douglas Parmee, London: Penguin Books, p. 251.

49 I say "or something very like it" because the view she rejects is one in which he is saying, among other things, that parents are morally permitted to love their children, whereas I take him more literally to be saying that parental love is what is most important in a case like this. (See below.)

the force of a natural attachment by giving it a roundabout and unconvincing justification from the impartial point of view."[50]

I find Annas' statement puzzling for a number of reasons. First, the idea would not be exactly that parents are permitted to love their children, but just what he says: parents' love for their children is more important [more important than (public) condemnation of immorality]. Permissibility enters in, but not in the way she suggests: it comes in because the implication is that because parental love is more important, it is okay to permit one's love for one's children to carry the day in a situation such as this. Annas says that this could not have much weight against their moral condemnation of Effi's actions. Why not? It seems clear that it would. Probably she says this because she was thinking of the mere fact that parents are permitted to love their children, not what Effi's father says, namely, that parental love is more important.

Second, why is there one thought too many? It is not as if *he* brought up the question of moral permissibility, an issue which, I would agree, it seems unnecessary to bring up, since we can hardly imagine that it would be *wrong* to welcome their daughter back (though we can well imagine that it would violate some community mores). Had he brought up the question of moral permissibility, I would agree that it is one thought too many; not, however, because one should not think about morality when one's son's or daughter's welfare is at stake, but simply because it should be obvious (at least to us, in our era) that what Effi's father proposes to do is not immoral. But in fact he is not raising the issue of moral permissibility. He is replying to his wife's worry that what they are doing is wrong, and is claiming that it is not wrong. It is hard to see how his thought could be one thought too many.

The third thing that I find puzzling about Annas' remark is this: how would this thought – the thought that dropping their moral condemnation and welcoming her back is not wrong – *weaken the force of a natural attachment?* I'll address this in a moment, but first I should acknowledge that there is another possible reading of Effi's father's remark.

50 Annas, 1984, p. 28. The details of her interpretation of Effi's father's remark are not important here, since I am citing her discussion only to have a better understanding of what counts as one thought too many, and what makes it "too many." But for those who are curious, here is her understanding of Effi's father's remark: "he realizes belatedly that it cannot be right to see morality as assimilating deep love and commitment to pathological weakness and self-indulgence" (p. 28).

I construed his remark to be that because parental love matters more, taking Effi back is not wrong. But it might be claimed that he is *not* speaking from a moral point of view but is rejecting it (and I do mean *it*, not merely rejecting the point of view of conventional morality). "Morality be damned! Parental love matters more." I don't think this is what he means, but I don't want to press that. I am interested in how things would be different if this were what he meant. Would there no longer be one thought too many? Perhaps, but that is not entirely clear. He would be rejecting the idea that moral justification is appropriate here; so if the one thought too many is the thought that what is being proposed is morally justified, that thought is absent. Notice, however, that he is saying that parental love matters more, *not* simply that it matters more to him, but that it matters more – presumably from an impartial perspective. If taking up that perspective involves one thought too many, then the reinterpretation of what he is saying would not take care of the problem. There would still be one thought too many. The thought would change, but its status as "one thought too many" would not.

The other question I want to ask is this: Would we think better of Effi's father if his remark were, or meant, "Morality be damned! Parental love matters more!"? Is it better to say that morality does not matter here – does not matter as much as parental love, anyway – than to say that morality endorses what they are doing? I don't think so. To take that view would seem (by conversational implicature) to endorse the notion that it actually *is* morally wrong of them to drop their moral condemnation of Effi for the adulterous affair that she had many years earlier. It seems much more apt to deny that it is morally wrong to do so, and to affirm that it is right to welcome her back into the family, and into their home.

Now, back to the question of why the alleged one thought too many – the thought that what they are about to do is morally right – would weaken the force of a natural attachment. (I am referring to Annas's claim that this "would be a classic case of…having 'one thought too many,' weakening the force of a natural attachment by giving it a roundabout and unconvincing justification from the impartial point of view.") If the thought is to have any effect on their love for Effi, it would most likely be to *strengthen* it, for they now will realize that they need not feel guilty or otherwise uncomfortable about welcoming their daughter – a "fallen woman" – back into the family. And they can quit fretting about what people will think; they know that what they are doing is right. So, why is the thought supposed to interfere, and weaken their love? Maybe it is felt

that what will dominate their thoughts and feelings is a sense of moral rectitude rather than love for their daughter. Maybe the worry is that they'll welcome her back *because* it is right, rather than because they love her. I don't see any reason why the thought that what they are doing is right would be likely to take over in this way. They welcome her back because they love her, and in the knowledge that what they are doing is right. Perhaps some would say that their love is mediated, and should be unmediated; the thought is one too many because it would be better, critics will say, if their love were not so qualified. It would be better if they could act from love without caring about, or paying any attention to, whether what they are doing is morally right.

There is some truth to the last point – but only because it seems to us so obvious that they were doing what is morally right that attention to the question of moral rightness seems silly. But it is not the case that in general, love should be unmediated by moral considerations. We are grateful that David Kaczynski did not think that love for his brother, or his relationship with his brother, was more important than going to the FBI with evidence that his brother was the Unabomber. We are glad that his love was not "unmediated." And this is not a case where we are simply glad because of the results. We admire him for his courage, and we do not judge him cold and unloving.[51] Of course it might be replied that happily, most of us do not have the Unabomber or any other mass murderer as a close relative, and therefore our love can and should be unmediated. I disagree, not because you never know for sure – your brother or sister may be a murderer too – but because these particularly dramatic moral considerations are not the only ones that merit our attention, and that our love for someone may tempt us to overlook. We need to be ready to assess the moral claim that the competing

51 In fact his love for his brother (evident in the lecture he gave at Indiana University, Nov. 11, 2005) is quite dazzling. He did everything he could to provide his brother with first-rate legal counsel and sought repeatedly to visit him, despite Ted's unwavering refusals; in fact he continues to write to him regularly, reiterating his love for him, despite Ted's refusal to have any contact with his family. David Kaczynski is evidence that one's love for another need not be unmediated for one to be committed to standing by the person – in some way – come what may. There are different ways of standing by someone, and some are fully compatible with recognizing and addressing the grave wrongs (and in this instance, crimes) that the other has committed (and may continue to commit if one does not intervene).

considerations make on us, and I see no reason to doubt that we can do so without it weakening or cheapening our love.[52]

In short, I don't see how the thought that what one is doing is morally right or permissible could be one thought too many unless it is so obvious that what one is doing is morally justified that there is no need to think about it.

VIII.

To recapitulate: After explaining why I think it more fruitful to assess the objections of virtue ethicists to Kantian ethics than to consider whether a Kantian virtue ethics is possible, I turned my attention to the (somewhat vague) objection that Kantian ethics is excessively impartialist, or overvalues impartiality. Focusing then on more precise statements of the objection, I evaluated the claim that Kantians cannot allow me to take the fact that someone is my good friend, or my child, or my sibling (or even just someone I like) as a reason for doing something for that person that I would not generally do for most others. I asked first whether the claim – the alleged fact – is correct, whatever its applicability as a criticism of Kant's ethics. I argued that with two important qualifications, it is correct. The first qualification is that it depends on what it is that we are doing. There are things that we ought to do for anyone, not only our friends and relations, and neither subjective preference nor special ties license us to ignore the needs of others altogether. The second (and closely related) qualification is that the fact that it is my friend or relative is a legitimate reason for doing *X* for him only if it is not impermissible to do *X* in the first place. With these qualifications in place, I claimed that while this position may be at odds with consequentialism, it is not in conflict with Kantian ethics.

Nonetheless, there might be something else in Kantian ethics that critics find unduly "impartialist," and I tried to ferret it out by utilizing Susan Wolf's novel characterization of impartialism. A disagreement that

52 I should clarify that it is not my claim that the fact that someone is a close friend or relative should never affect our judgment about how to deal with a particular (suspected) wrongdoing, and specifically, about whether to turn the person in to the authorities. Our relationship with someone does count in the moral balance, particularly when the wrong was (morally) minor – a harmless wrongdoing or perhaps not even a wrongdoing at all, apart from being a violation of the law – and all the more so if in addition the person is likely to be punished too harshly.

critics have, I suggested, concerns how ready we should be to take up a detached perspective towards those we love or towards our relationships with them. This is a genuine disagreement, but Kantians have no cause for embarrassment, or for worry that Kantian agents have one thought too many.[53]

53 Earlier versions of this paper were presented at a conference on ethics and impartiality at the University of Utah, at a UK Kant Society conference (Keele, England), at McGill University (as a colloquium cosponsored by the philosophy departments of the University of Montreal and McGill University), Auckland University, Indiana University, and the University of Kentucky. I am grateful to discussants, and to Wade Robison, who presented a comment at the Utah conference. I would also like to thank Justin Brown, Sarah Conly, Michael Gorr, Mark Jenkins, Samuel Kerstein, Henry Richardson, David Suryk, Christine Swanton, and David Velleman for their helpful comments.

Kant on Solving Moral Conflicts

Andrea Esser

Moral conflicts and dilemmas are major issues of contemporary debates in moral theory. But, much to our surprise, we find that in Kant's ethics – which is still one of the prominent positions in Moral Philosophy – there is no *problem* of moral conflicts. The law of practical reason, the Categorical Imperative, is meant to provide us with moral guidance for all possible cases of human practice. Kant is convinced that this law, therefore, also guides us in situations, which, at first sight, confront us with conflicting claims. From his point of view, such conflicts can easily be resolved in that the Categorical Imperative either proves one of the reasons of obligation not to be binding, or simply requires us to follow the stronger one. Therefore, Kant suggests that practical law enables us to identify, in any situation, what we are obliged to and what not, and act on it accordingly. This sober view of situations leaving those involved tormented by awful feelings long after the event, has often been criticized as radical and rigoristic. The problem of moral conflicts clearly shows, according to critics, how inhumane the outcome of a merely principle-based moral theory is when applied in practice.[1]

One can only come up with this judgment on Kant's ethics, if confining it to what is set out in the "grounding works" (the "Groundwork of the Metaphysics of Morals" (1785) and the "Critique of Practical Reason" (1788)) without considering the concrete applications presented in "The Metaphysical First Principles of the Doctrine of Virtue" of the "Metaphysics of Morals" (1797).[2] According to the Kantian self-conception, however, the early grounding works by no means represent his entire ethics. Rather, they only serve for the preparation of a so-called metaphysics of morals, the task of which is to

1 See Bernard Williams, 1973, *Problems of the Self*, Cambridge: Cambridge University Press; and Ruth Barcon Marcus, 1996, "More about Moral Dilemmas", in: H. E. Mason (ed.), *Moral Dilemmas and Moral Theory*, Oxford: Oxford University Press, pp. 23–36, at p. 24.

2 As to the possible causes of this development cf. Alwin Diemer, 1953/1954, "Zum Problem des Materialen in der Ethik Kants", in: *Kant-Studien 45*, pp. 21–32.

devise the contents of the general moral principle and therefore show consideration for the special nature of man and the particular circumstances given in each individual case. If also including the reflections made in the "Doctrine of Virtue" in one's discussion of moral conflicts, it turns out that Kant's ethics does not at all decide concrete moral issues by means of abstract principles only.

In order to solve particular moral problems, I will demonstrate in Part I of this essay, that Kant's "Doctrine of Virtue" calls on us to make moral deliberations which are governed by formal principles but remain open as to the concrete action taken. Correspondingly, moral conflicts, and the feelings emerging therein, are reflected upon and coped with in such a process of deliberation. For the sake of demonstrating how the current problem of moral conflicts can be dealt with, within a Kantian approach, I refer to the example of 'lying' in Part II. This example is chosen by Kant himself, in his short essay "On a Supposed Right to Lie because of Philanthropic Concerns". The case outlined therein is often referred to in contemporary debate to demonstrate the impracticability of Kantian ethics and its underlying principle. Using this example, I will present some misunderstandings in interpretation and also show how a satisfactory solution of this case can easily be achieved without having to relinquish valuable insights of Kant's ethics. For this purpose it is not only important to take into account the difference between the juridical and ethical perspectives, but also determine the Kantian concept of the Maxim adequately: as a hypothesis of the basic moral attitude of one's own acting, or that of others, and not as psychological motive. This element is dealt with in Part III. In my summary, Part IV, I propose firstly: general principles in Kantian ethics are not entirely separate from a pragmatic dimension and secondly: not every practical problem constitutes a moral problem.

I. The Problem of Moral Conflicts

I.1. The contemporary view

Telling lies to others, holding them in contempt, or denying them one's help is not morally desirable. Almost certainly we would criticize somebody for referring to such behaviour as morally permitted or even required. The confidence we have in our judgements in view of these general questions is shaken, when faced with concrete cases. Often the

circumstances are so complex that we find it hard to come up with a clear moral judgment nor can we tell which action would be morally correct in this case: should we tell someone our honest opinion, if we know this will hurt them deeply, or if this would be of no use to the person and only cause them harm? Might it even be necessary to deny our help to someone, in order for them to solve their problem independently? Is there any kind of behaviour that is subservient, and cowardly enough, to really deserve our contempt, after all? What needs to be done if a morally motivated decision would detract from the happiness of others or even, cause them suffering? How are we to act in a situation as proposed in Kant's short essay "On a Supposed Right to Lie because of Philanthropic Concerns", in which a murderer asks us to forsake a friend and yet the act of saving our friend would require us to lie, and betray our feeling that we ought to stay honest?

In all these cases, *moral* considerations are in conflict with each other and seem to make contradictory demands on us. The contemporary discussion in moral philosophy refers to such situations as "moral conflict". Their structure is generally described as involving a person who *ought* to do *x* as well as *y*, but *can*not do both.[3] Conflicting normative claims, each taken individually, strike the agent as justified and urge towards realisation. Although it is generally assumed that a moral "ought" also suggests a "can" – as the principle "ought implies can"[4] demands – we seem unable to cope with the moral claim made on us in such cases. Curiously, Kant's ethics deals extremely briefly with the problem of "hard cases" and is also unconcerned with the ensuing emotional consequences.[5] For this reason Kant's ethics has often been reproached for adhering rigidly to its principles and so failing to make any useful contribution to the issue of "moral conflicts".

3 Cf. Williams, 1973, p. 121; Ruth Barcon Marcus, 1980, "Moral Dilemmas and Consistency", in: *The Journal of Philosophy* 77, pp. 121–136.

4 A passage, which clearly demonstrates that this principle is not a Kantian one [*MM* 6: 380.

5 Cf. Monika Betzler, 2004, "Sources of Practical Conflicts and Reasons for Regret", in: Peter Baumann/Monika Betzler (eds.), *Practical Conflicts: New Philosophical Essays*, Cambridge: Cambridge University Press, pp. 197–223, at p. 199.

I.2. Kant's position in the "Metaphysics of Morals"

Turning to one of the few passages in which Kant addresses the problem of moral conflicts, we find that not only does he pay little attention to this issue, but he even rules out any possibility of colliding duties and obligations altogether:

> A *conflict of duties (collisio officiorum s. obligationum)* would be a relation between them in which one would cancel the other (wholly or in part). But since duty and obligation are concepts that express the objective practical *necessity* of certain actions and two rules opposed to each other cannot be necessary at the same time, if it is a duty to act in accordance with one rule, to act in accordance with the opposite rule is not a duty but even contrary to duty; so a *collision of duties* and obligations is inconceivable (*obligations non colliduntur*). However, a subject may have, in a rule he prescribes to himself, two *grounds* of obligation (*rationes obligandi*), one or the other of which is not sufficient to put him under obligation (*rationes obligandi non obligantes*), so that one of them is not a duty. When two such grounds conflict with each other, practical philosophy says it is not that stronger obligation takes precedence (*fortior obligatio vincit*) but that the stronger *ground of obligation* prevails (*fortior obligandi ratio vincit*) [*MM* 6: 224].

Obviously Kant reserves the terms "duty" and "obligation" for such claims which are "objectively necessary" in the practical sense. Whether or not we are obliged to do something specific, according to Kant, is neither governed by the circumstances or the consequences of an action, nor by our emotional attachments or inclinations, but by moral law alone: the Categorical Imperative. It tells us "what obligation is" [*MM* 6: 225]. Consequently, insofar as there is only this *single* source of duty and obligation, as Kant thinks, there cannot be any collision in this respect at all. At most, the "grounds of obligation" (*rationes obligandi*) [*MM* 6: 224], or the "motives" (as it reads in a different passage), may come into conflict with each other. But in such cases, too, Kant invariably holds a conflict to be resolvable: By means of the Categorical Imperative, either one of these grounds can be proven to be not actually obliging (to be only "prima facie" reason), or else "the *stronger* ground of obligation" simply "prevails" (*fortior obligandi ratio vincit*) cf. [*MM* 6: 224]. However, not all readers of Kant's ethics find these remarks convincing. In order to demonstrate how unacceptable Kant's scheme for a solution of this problem is, reference is often made to his short essay "On a Supposed Right to Lie because of Philanthropic Concerns". Discussing Benjamin Constant's position in this essay, Kant deals with the question: whether an unconditional obligation to truthfulness exists and if this obligation is

to be fulfilled irrespective of the intentions, objectives and moral qualification of the other person. Whether, for instance, we have to tell the truth even to a murderer asking for a friend hiding in our house.

By the common definition of a moral conflict one could comprehend the situation as a conflict between two obligations, according to which we ought to help the friend, but also should tell the truth. Now, if the murderer forces a response from us, we cannot do both and are thus faced with two irreconcilable moral claims. Kant's solution seems terrifyingly simple: We must tell the truth – as demanded by the moral law – and must do so entirely regardless of the consequences and our close attachment to the friend.

I.3. Contemporary criticism on Kant and Kantian ethics

It is no wonder that this conclusion strikes many as radical and inhumane.[6] The application to a concrete case seems to clearly reveal the absurdities a single-principle-based moral theory can lead to. The unconditional nature of moral obligations in Kant's ethics, according to critics, also sacrifices the actual human dimension of our acting and our moral feelings for the sake of the "purity" of the principle.[7] Don't we hold a person who acts on principles only to be morally deficient rather than a moral model to be followed? Isn't there an unbearable tension between the moral law and our moral intuitions and feelings?

Many present-day authors arrive at the opinion that situations of moral conflict cannot, at all, be adequately resolved by means of Kantian ethics.[8] Conflicting values, they reproach the Kantian conception with,

---

6 Cf. Hariolf Oberer, 1986, "Zur Vor- und Nachgeschichte der Lehre Kants vom Recht der Lüge", in: Georg Geismann/Hariolf Oberer (eds.), *Kant und das Recht der Lüge*, Würzburg: Königshausen und Neumann, pp. 7–23, at p. 7; cf. Simone Dietz, 2002, *Der Wert der Lüge*, Paderborn: Mentis Verlag, p. 139; cf. Jens Timmermann, 2000, "Kant und die Lüge aus Pflicht", in: *Philosophisches Jahrbuch 107. Jahrg.* II, pp. 267–283; cf. Ottfried Höffe, 1990, *Kategorische Rechtsprinzipien. Ein Kontrapunkt der Moderne*, Frankfurt am Main: Suhrkamp Verlag, p. 195; cf. Oswald Schwemmer, 1973, "Vernunft und Moral", in: Gerold Prauss (ed.), *Kant. Zur Deutung seiner Theorie von Erkennen und Handeln*, Köln: Kiepenheuer & Witsch, pp. 255–273, at p. 255.

7 Bernard Williams, 1994, *Ethics and the Limits of Philosophy*, Cambridge/Massachusetts: Harvard University Press, p. 248.

8 "Kant imagines that he had provided a single-principled framework from which all *Maxims* flowed. But Kantian ethics is notably deficient in coping with

cannot be sorted according to a common currency. Bernard Williams argues that in order to achieve a "sound and humane" ethics, one must not proceed from a single principle but assume a *plurality* of values and obligations. Also, one needs to consider that a moral conflict resembles a conflict of desires and feelings than one of rational convictions. As a consequence the claim a value makes on the agent does not simply disintegrate once a decision has been made, but continues to exist – like a desire, which remains unfulfilled. It is for this reason that even having solved a situation of moral conflict, "all-things-considered", we do, still, feel regret. So, this feeling of regret is held to indicate that – even after the conflict as a matter of fact has been resolved – the unrealised normative claim does not cease to be a reason for action and thus continues to remain in existence. Firstly, Williams claims to have shown by this deliberation, that the so-called "single-principle-architectonic" has to be abandoned and, secondly, to have furnished a rational and moral explanation for the feeling of regret. At the same time he believes to have solved a problem of Kant's ethics. In his view, it entirely fails to explain such feelings[9], or adequately come to terms with moral conflicts.

One only arrives at this view, though, by misunderstanding Kant's ethics as a theory that is rigidly modelled on abstract principles and by reducing the moral evaluation to a subsumptive mechanism. On this basis, I present an understanding of Kant's ethics, which provides both a satisfactory and humane solution of moral conflicts as well as a rational explanation of the feelings involved therein. This reading of Kant's ethics is not only critically directed against William's explanation of moral regret, but further seeks to clarify some misunderstandings of central Kantian thoughts that have emerged particularly in the more recent discussion as well.

---

dilemmas. Kant seems to claim that they don't really arise, and we are provided with no moral grounds for their resolution", Ruth Barcon Marcus, 1987, "Moral Dilemmas and Consistency", in: Christopher W. Gowans (ed.), *Moral Dilemmas*, Oxford: Oxford University Press, p. 192.

9 Cf. Rüdiger Bittner, 1992, "Maximen", in: G. Funke (ed.): *Akten des 4. Internationalen Kant-Kongresses Mainz*, Teil II,2: Sektionen, Berlin /New York: De Gruyter, pp. 485–498.

I.4. What does *Government by Principles* in a Kantian ethics mean?

In his "Groundwork of the Metaphysics of Morals" Kant answers the question of whether there can be a binding moral guidance for all rational beings, not by a material law, but by a formal method of determining the will through practical reason.[10] He points this out at the very beginning of this work: All kinds of goods and values generally held in high esteem could be specified for determining the concept of the good. However, it is possible that all these goods and values each lose their positive-moral quality under concrete circumstances and turn into evil. They are not good "without limitation", rather, they are good only in relative terms, i.e. in relation to morally good ends and thus only on certain conditions. As a consequence of these considerations Kant refuses to start from material values in his grounding of the moral law. Correspondingly, in Kantian ethics – as though it may come as a surprise – not even "life" is seen as a "value in itself" which is to be protected under any circumstances, and by all means.[11] There are circumstances conceivable under which it is not morally desirable to preserve life. This may be the case, if one's own life can only be saved at the cost of the lives of others. Or if preserving one's life, as Socrates explains at the end of the Platonic dialog *Phaidon*, this only means a minor extension of its duration in time, but its quality is beyond our control. Similar deliberations may, of course, be made with reference to all other goods such as: health; wealth; courage; and intelligence – even if we, at least in our everyday use of language, doubtlessly hold them in high esteem. All of these values, according to the Kantian thought, only gain moral value, if employed for morally justifiable or demanded ends. As they themselves cannot guarantee their correct use, though, they are no more than means to ends whose moral quality still is to be examined. For this reason Kant does not believe that the question of what is to be named morally good can be satisfactorily answered by simply enumerating generally approved values. The actual

10 Cf. [*G* 4: 427].

11 "And Sokrates said: "Crito, those whom you mention are right in doing as they do, for they think they may gain by it; and I shall be right in not doing as they do; for I think I should gain nothing by taking the poison a little later. I should only make myself ridiculous in my own eyes if I clung to life and spared it, when there is no more profit in it.", Platon, "Phaidon", 1960, in: Harold N. Fowler (ed.), *The Loeb Classical Library Vol. 1*, Cambridge/Massachusetts: Harvard University Press, 116e7–117a3.

challenge of moral philosophy in Kant's eyes is to come up with a general criterion of examination.

In the "Groundwork of the Metaphysics of Morals" therefore a method is developed which is meant to express the idea of an "unconditional moral end". As supreme restrictive condition of the will, it represents the general criterion for the justification of any moral claim whatsoever. Various formulations of the Categorical Imperative put this thought into words. The formal nature of the moral law and its various formulations, one could say, are its entire substance.[12] Taken on its own, this law does not give any recommendations or lead straight to specific contents. The formulae of the Categorical Imperative merely explicate the rational conditions of morality and thus are criteria of moral obligation in general. Therefore the Kantian ethics neither asks for the realisation of abstract moral contents, nor does it (regardless of the concrete circumstances) demand that a *specific action* be performed. Kantian ethics, as we will see, does not directly deal with the subject matter of actions, anyway. Rather, actions are always subjects of the ethical judging as manifestations of determinations of the will only.

I.5. The will as subject matter of practical philosophy

Generally speaking only what "is represented as possible (or necessary) through a will"[13] comes within the realm of practical philosophy. This, in any case, is the differentiation Kant makes between practical and theoretical philosophy in the "Introduction" to his "Critique of Judgment". "Possible through a will", of course, in this context means: possible through a will that is capable of self-determination and insofar free. It is only our will we are in control of at any time, and under all circumstances, and which we can determine autonomously, i.e. through

12 Cf. Georg Anderson, 1921, "Die 'Materie' in Kants Tugendlehre und der Formalismus der kritischen Ethik", in: *Kant-Studien 26*, pp. 289–311, at p. 302. Equally Thomas Pogge who concedes purity only at the level of the supreme moral principle, not, however, on the level of special moral precepts and therefore conceives Kant's examples as tests of *Maxims* in the human world, cf. Thomas Pogge, 2000, "The Categorical Imperative and the Natural Law Formula", in: Otfried Höffe (ed.), *Grundlegung zur Metaphysik der Sitten. Ein kooperativer Kommentar*, Frankfurt am Main: Vittorio Klostermann Verlag, pp. 172–194, at p. 181.

13 Immanuel Kant, "Kritik der Urtheilskraft" [5: 172].

the moral law. It is true that it needs to be taken into account that Kant does not regard the determination of the will as detachable from its realisation, as has often been alleged. But it is only in the "Metaphysics of Morals" that this connection becomes really evident. The terminological modification made therein rules out any confusion between the "pure will" (which in the "Metaphysics of Morals" is stated to be "practical reason itself" [*MM* 6: 213]) and the "empirical will", a confusion, which is still possible in his early writings. This modification clearly identifies the latter as "capacity for choice" (Willkür). Thus it is a particular case of the "capacity for desire" (Begehrungsvermögen)[14], namely the capacity for desiring in accordance with concepts, ("insofar as it is joined with one's consciousness of the capacity to *bring about* its object by one's action" [*MM* 6: 213]. This understanding conceives the will as an intention to realise and differentiates it strictly from mere wishing, which does not make any effort towards its realisation.[15] So, if the general formula of the Categorical Imperative[16] calls for *action*, this is not in contradiction to the fact that the Categorical Imperative does not choose the action itself but its underlying determination of choice (Willkür) as its subject matter. The intention to act in Kant's understanding is already comprised in the concept of choice (Willkür) and thus the normative orientation is not directed at desires or yearnings, but at the determination of choice (Willensbestimmung) as a self-standardization urging for realisation.[17]

I.6. Can the "Good Will" come into conflict with itself?

Which determination of choice (Willensbestimmung) is to be referred to as good, though, we cannot be told by any set of generally accepted values or socially desired duties. Which ends we should pursue, what we are "bound" to in a moral sense in the first place, and what should be considered a duty and a value in each case is stated by the moral law alone. This law, on the one hand, subjects all our volition to the negative condition: A morally permitted end is only one which can be realised in

14 [*MM* 6: 2].

15 [*G* 4: 394]. As to a differentiation of "will" and "wish" also cf. [5: 177]; and the introduction to the "Metaphysik der Sitten" [*MM* 6: 213].

16 Cf. [*G* 4: 421].

17 A person who for instance merely harbours immoral fantasies, but does not let these become effective through acting, proves moral precisely in that his or her acting is orientated by morally binding principles and not by these fantasies.

an action, whose *Maxim* is conceivable and desirable as a general law. On the other hand, in the "Doctrine of Virtue" the moral law names certain material ends which the agent should rate higher than his private ends; these are not specific contents of action, but "ground", as Kant puts it, a law for the *Maxims* of the agent. This means: the agent should form *Maxims* that correspond to these objective ends. So Kant's "Doctrine of Virtue" does not content itself with providing a merely negative formal criterion for refuting egocentric *Maxims*, which cannot be generalised. It vests the general ethical orientation with a positive productivity: only this action accordingly realises a binding "value", the *Maxim* of which fulfils the demand that man be an end in itself and capable of autonomy.

So the moral law is not a rigid principle, but governs a process of reflection determining which *Maxim* in the particular situation given is morally justifiable and which end is to be considered a "value" in the first place. Consequently, in Kantian ethics there "are" no set values, but these "come into being" only through applying the Categorical Imperative to concrete cases. The Categorical Imperative is neither a value in itself, nor does it yield abstract values independently of concrete circumstances. It is the criterion of what *ought to* represent a value for choice (Willkür), in each case, under particular circumstances. It "as such only affirms what obligation is" [*MM* 6: 225].

So, if we see ourselves in a conflict between values, this is – as Williams can be understood – because we consider general goods desirable irrespective of the concrete circumstances and think we ought to realise them. Seen from this general viewpoint, there are many values, in any situation of our lives, whose realisation is to be aspired to. And generally speaking like this, a morally acting agent should seek to choose the *right* ones from this "set" of general values and obligations. Finding ourselves in a particular situation, though, it soon turns out that, on the one hand, we are unable to realise all values. On the other hand, often we are not sure, anyway, which one of these abstract values actually is to be realised in a given situation, i.e. which one of these abstract values in this concrete situation remains a value at all. In order to help us to progress in this question, Kantian ethics provides us with a criterion which releases us from the absolute claim of abstract values by indicating which *Maxim* is permitted and called for *under these particular circumstances*, so that it is of moral "value". A conflict between values from a Kantian perspective, therefore, only consists in the abstraction of a thought experiment. In the concrete situation this conflict vanishes with the application of the Categorical Imperative. For its very *formal and situational nature*,

however, it rules out that we *ought* to do something we *can*not do. A moral conflict between values consequently is impossible.

I.7. Kant's rational explanation of regret

In a difficult situation in which we believe ourselves to be faced with contradictory normative claims, according to the initially quoted passage of the "Metaphysics of Morals", the "stronger reason" of obligation keeps the upper hand. The other claim is dismissed in the moral reflection as non-binding. Should we, in a particular case, only perform the moral reflection in retrospect thereby realising that we have acted for immoral *Maxims*, then, of course, we may regret this. The feeling of regret in this case has a moral reason and, therefore, is a moral feeling. The rationality of the preceding decision is called into question by this result. However, if the moral reflection proves that we have acted on the basis of morally correct *Maxims*, then, according to Kant, our regret cannot actually be for the particular decision we made in this case, or for the resulting actions. Even if in such a situation we have made a terrible (but in the terms of the Categorical Imperative) yet *rational* decision, no *moral* reason is given that continues to call for being acted upon, nor does the fact that it remains unfulfilled need to be regretted. If one namely claimed this to be the case, one would not only come into contradiction with the requirements of the Categorical Imperative, but also have to revoke the idea that there exist rational reasons of moral obligation in the first place. This is because it is not our feeling but the moral reflection, which tells us what we are bound to in a *moral* – not a psychological – sense.

The regret felt by many people after having experienced so-called hard cases, may be indicative of the fact that the agent is a decent, responsible person. Despite having acted to the best of their knowledge and belief, they still wish their actions had produced a happier outcome for everyone affected. This regret is not irrational and, within the Kantian ethics, does not appear as such either. After all, it is absolutely rational to feel regret about having lost something we value. Our regret may also be for the awful circumstances, the tragic consequences, or our own helplessness in this particular case. It may be felt by someone observing, or involved in, a tragic situation, which, in turn, may prompt one to compensate for a harm; to express one's sympathy, or appreciation; or to apologize for the consequences of an inevitable breach of obligation. From this perspective regret is an emotional response for which there are

rational explanations also within a Kantian ethics[18], even if no moral reasons are assumed.

However, inferring an unfulfilled obligation from a factual feeling – a conclusion William draws – is always "endlessly hypothetical".[19] Kant has set this out in his "Critique of Judgment" for the feeling of pleasure brought about by the aesthetic judgment. It is true that on the basis of such a conclusion much may be speculated about how some feelings actually evolve. But which explanation out of so many is the "correct" one for explaining the occurrence of a feeling in each particular case? It is hard to tell which might help qualify the occurring feeling as a "moral" one. Whether or not a factual feeling of regret is to be referred to as "moral", is neither decided by its persistence nor by the context in which it occurs. For a feeling to be rated as moral it is required that it also has a moral reason.[20] This also means that a regret lingering for a long time is not necessarily indicative of a claim which has been omitted, but still is morally, correctly called for. Simply assuming so is based on uncritical concluding. One starts from given facts and assumes something "behind them" without examining this assumption itself.

If Kant's ethics, according to this reading, does not subject us to the dictates of a general principle, after all, but even manages to free us from unjustified claims and tormenting feelings, one is right in expecting this to be equally true for the resolution of concrete moral conflicts. Looking at the most prominent example from the so-called "Essay on Lying", its proposed solution does not concur with the reading of Kant's ethics presented here. Rather, we gain the impression that Kant sees an absolute value in truthfulness, which is not relativised in the slightest by the given circumstances and, therefore, doubtlessly is to be ranked above other values such as life, philanthropy and friendship. It is a widely-held view that Kant's ethics resolves moral conflicts in the rigid way mentioned. On closer inspection of this essay, however, one encounters quite a few deliberations that can shake this opinion.

18 See Betzler, 2004, p. 215 f.

19 Immanuel Kant, "Erste Einleitung in die Kritik der Urtheilskraft" [20: 238].

20 For the case of compassion for instance this would require that one is really touched by the suffering of the other person and not only shocked by the thought of meeting a similar fate oneself.

II. The Classic Example of a Moral Conflict: Lying Because of Philanthropic Concerns

II.1. Is there a "Right to Lie because of philanthropic concerns"?

In the discussion about the problem of moral conflicts, reference is often made to Kant's short essay "On a Supposed Right to Lie because of Philanthropic Concerns". However, this famous example is not particularly suitable for a discussion of moral conflicts.[21] Its subject matter is not really a *moral* problem, but one within the philosophy of law. The question the essay deals with is not: "Should I save the friend or should I tell the truth?" but "Is there a *right* to lie?" The question of whether one has a right to lie, or if lying can even be legitimised as a duty of right, is a problem not solely of the cited example. Kant stresses the basic significance of this question and, therefore, discusses it primarily in fundamental terms. Constant gives an affirmative answer to this question. Kant, by contrast, differentiates two aspects: as far as the agent is concerned, he concedes that the person unduly compelled to make a statement does *not* do any wrong if he or she is untruthful.[22] So, by no means does Kant, in this essay, require the agent to tell the truth. But there is another, more general aspect on which Kant places emphasis: The concession at the level of the individual has not yet decided the basic problem posed by Constant: namely whether there can be a *general right* to untruthfulness. Kant negates only this point arguing that neither a general right, nor an obligation to lie, can be consistently derived. Truthfulness, according to Kant, is an absolute obligation. But this is not because we need to be able to rely on the correct use of language for sealing contracts and conducting legal transactions. The call for truthfulness, as a matter of principle, guarantees the possibility of declarations in the first place. Truthfulness is a prerequisite for the contract-based liability also of any particular legal system. This consideration marks the

21 Cf. Hans Wagner, 1986, "Kant gegen 'ein vermeintes Recht, aus Menschenliebe zu lügen'", in: Georg Geismann / Hariolf Oberer (eds.), *Kant und das Recht der Lüge*, Würzburg: Königshausen und Neumann, pp. 95–103; cf. Julius Ebbinghaus, 1954, "Kant's Ableitung des Verbots der Lüge aus dem Rechte der Menschheit", in: Georg Geismann / Hariolf Oberer (eds.), *Kant und das Recht der Lüg*e, Würzburg: Königshausen und Neumann, pp. 75–85, at pp. 77, 83.; cf. critically Timmermann, 2000, p. 271.

22 "He himself, properly speaking, does not *do* harm to him who suffers thereby; but this harm is *caused* by accident" [8: 428].

other aspect and the level at which Kant discusses the issue: it is not a question of whether (and under which circumstances) an individual actually ought to tell the truth. What is at stake is whether a *general* obligation to truthfulness exists, or whether this is the object of a particular legislation. Insofar as the truthfulness of declarations is not an arbitrary subject matter, but the basic condition of the possibility of contracts, there cannot be any exception to it. The same applies to exceptions to principles: there cannot be any, because otherwise they no longer would be principles but references to further principles which, in turn, determine and regulate the case of exception.[23]

So Kant negotiates the question of an obligation to truthfulness at the level of natural law (the system of binding primary precepts governing our acting conceived as independent of any concrete legal system). In other writings Kant has sufficiently set out that concrete deliberations must not play a part in this connection. If, finally, the empirical subjectivity decided situationally which principles apply, one could, of course, do without any reasoning and leave it up to the judgment of the individual. But which criteria should he or she go by? In order to avoid this gap, Kant objects that hardship and difficulties that may be engendered in the application and execution of principles (for instance by coincidence, or through their undue use) may call the validity of these principles into question. At the level of principles, 'right' as the system of obligation in general, cannot do without the demand for truthfulness of the declarations. This means that Kant, in this matter, has not let moral psychological considerations entice him, or tried to prevent a "landslide"[24], but has argued systematically.

However, on the conditions of the reality of concrete cases, special laws have to be drafted, which see to it that the *subjective truthfulness* (not as a principle) of the agent, in each case, is protected against a misuse for wrong ends. This problem raises the question whether there is any legal condition on which one can be released from the general obligation to truthfulness. This issue is decided by politics. By no means can one proceed independently of the demands of the general principles of law.

If it is a question of finding out which *legally* acceptable option for acting, with which moral value, the individual has in such situations, reference can be made to the "Right of Necessity" (Nothrecht)[25] which

23 Cf. [*G* 4: 424].
24 Cf. Timmermann, 2000, p. 273.
25 [*MM* 6: 235]; cf. Wagner, 1986, p. 95.

Kant concedes. The "Right of Necessity" represents a legal means for preventing the rigorism a "Strict Right" might assume at the level of the individual. It becomes effective in situations in which the legal system fails to enforce a correct action, because this action would be at least equally as threatening to the agent as the sanctioning of its omission. The right of necessity does not render ignoring a command – in the present case: false testimony – conforming to law, even less so turns it into an enforceable right. But it does become an action which no longer is a punishable offence. The right of necessity is not actually a right; it merely leads to an authorization and consequently bears a wrong name. Kant therefore refers to it as a merely "alleged right".

The ruling law in Germany and Europe at the level of concrete legislation protects the subjective truthfulness. However, it does not do so by permitting "lying" in general. Against the backdrop of a general call for truthfulness, it is exactly determined which statements, on the one hand, are legally relevant and then – measured by this value – need to be rated as "wilful deceit" in the sense of § 121 German Civil Code. This thought can be elucidated by looking at a case which has also been discussed within European Law: If, in a job interview, the applicant answers 'no' to a question about a possible pregnancy, even though it could be proven that she knew at the time that she was pregnant, although, colloquially speaking, she told a lie, is the contract void for this reason? § 611a of the German Civil Code (prohibition of discrimination) and the Council Directive 76/207/EEC of the European Parliament executed therein, prohibit the discrimination against a pregnant job applicant and, thus, also the question for an existing pregnancy. The Federal Labour Court therefore dismissed the legal action of the employee arguing as follows: First of all, the very question is inadmissible (according to § 611a). Secondly, it did not have to be answered truthfully, either. So, in total, no "wilful deceit" is given and the work contract remains valid. This case is in continuity of Kantian thinking insofar as here, too, the obligation to truthfulness as a principle is not questioned, but a subordinate rule is drafted, in order to avoid that the validity of the obligation of the normative system is shaken by factual wrongdoings. What is remarkable about this is the clarity of language: it is not "wilful deceit" which is then permitted, or assessed, as excusable under these specific circumstances. It is precisely that *no* wilful deceit has occurred. The Kantian line of thought would have to be accentuated in this way and the insights gained also adequately formulated: There cannot be a "general right to lie", but in some situations – which, however, are to be

carefully specified by the ruling law – one is not obliged to make true statements. Therefore, the statement made in such situations must not be referred to as a "lie", either.

II.2. Is it morally forbidden to ever tell a lie?

The strict call for truthfulness and the prohibition of lying, of course, can also be discussed from a moral-ethical angle. In analogy with the law case here, too, first of all one has to accept that a general obligation to, and permission of, untruthfulness – just like a general rule of exception from the obligation to truthfulness – cannot be conceived without contradiction. In any case these are the familiar results of Kant's earlier writings. By contrast, in the "Metaphysics of Morals", the general principles applying to all rational beings are examined for their human conditions of realisation and discussed with reference to concrete cases of decision. This late writing, therefore, is the correct point of reference for an investigation into how situations of moral conflict are to be discussed within Kant's ethics.

In the passage about lying in the casuistry of the "Doctrine of Virtue" – entirely in continuity of the "Groundwork of the Metaphysics of Morals" – truthfulness is characterised as a "perfect obligation to oneself". This passage suggests that Kant prohibits the negation of truthfulness, namely lying, and thus simply transfers the argumentation of the example from the field of law to that of ethics. Thus the rigorism Kant is criticised for obviously does not prove true for his philosophy of rights, but does so for his ethics. However, the same section of the "Doctrine of Virtue", besides the general prohibition of lying, also contains several examples of situations of moral conflict. Kant does not decide these situations by giving a concrete instruction as how to act. Also, considering the demand that ethics and duties of virtue are to be directed at the agent's *Maxims of will*, the impression of rigorism should, as I demonstrate in the following, vanish in the moral context.

To start with, it is important to remember that lying is not limited to language but can equally be realised in any other form of communication. There are all sorts of media of lying: a formulation, which cannot be proven to deliberately employ untruthfulness, a gesture, even a certain kind of behaviour. Not the external form, but the *Maxim* of the action marks the boundary between the overtly expressed unwillingness to reveal one's thoughts and telling a lie. According to Kant the *Maxim* of

intending to deceive someone, alone, qualifies as lying. The fact that the *Maxim* of deceit is also conceived as a breach of a "duty to oneself" is explained by the idea of what lying is all about, irrespective of its effect, namely the *actual* deceit of others. As such, it primarily concerns the agent himself. Lying, then, is a violation of "the dignity of humanity in his own person" [*MM* 6: 429], if the person who lies is not prepared to take responsibility for foreseeable, or likely, effects of truth. This also applies, if the agent assesses the consequences of truthfulness in this particular case as too unpleasant for him and justifies his decision against truthfulness by the special circumstances, or allows these circumstances to dictate his decision. But if this reflection on the *Maxim* yields that one of these deliberations is expressed in it, it is not an obligation, which collides with the moral call for truthfulness, but merely an interest. This may consist in refraining from a conflict with, or in avoiding upsetting, another person. In any case, though, an interest leads to the intention to not take responsibility for lying but to blame it on the situation. In doing so, the agent also deceives himself about his true *Maxim* when justifying his lie. So the "external lie" inevitably involves an "inner lie". Therefore lying deliberately negates everyone's capacity of freely setting ends for oneself, but also that of critical self-elucidation.

This thought is also backed by an example from the "Doctrine of Virtue" which is similar to the case presented in the essay on lying, but clearly a moral-ethical example: Kant drafts the scenario of a servant who is told to pretend to a visitor that his master is not at home. This way of acting on the part of the servant does not exactly trigger a crime committed subsequently by the visitor, but does not help to prevent it, either, even though this would have been possible, if the servant had acted contrary to the instruction. It is true that the servant does not share the guilt for what has happened in terms of *law*, but Kant does consider a responsibility "in accordance with ethical principles" [*MM* 6: 431]. With this example Kant confirms the validity of the prohibition of lying in the realm of ethics. He is of the opinion that the servant had to blame himself for the breach of a moral obligation. Still the basic parameters in this example differ from those of the law case: the servant is not faced with a life-threatening situation, but merely with the instruction of his master. Whatever may be the *Maxim* of the servant – whether he acts for absolute obedience to his master or out of indifference – in any case he has given preference to this *Maxim* over truthfulness and has acted accordingly, even though his *Maxim* was not a moral obligation. Examining the relationship between the servant's *Maxim* (which is not further explicated

in the example) and the ethical demand of the prohibition of lying, it is evident that no collision of obligations occurs, but the moral obligation only contravenes a mere interest. In acting as he does, the servant has made a decision in favour of the realisation of his interest and against the obligation. By his action, he breaches his obligation entirely regardless of the ensuing consequences. In this way Kant sustains the prohibition of lying also in the moral context; as no moral obligations are conflicting, this is not an example of *moral* dilemma, either. Moreover, one has to draw the conclusion that a rigorism, in this context, is only engendered if moral considerations are directly related to actions and to *Maxims* of actions – as demanded by Kant.

So Kant's ethics provides a simple and clear-cut solution for moral conflicts in the context of the prohibition of lying: the ethical *prohibition* of lying – as does all other ethical legislation – does not immediately *command* specific actions. As *ethical* principle, its subject matter is the *Maxim* of the agent. *Maxims* generally grant the agent some scope in his decision about which concrete actions they can be realised by in each case. Consequently, Kant states what is to be done in the concrete case:

> judgement can decide what is to be done only in accordance with rules of prudence (pragmatic rules), not in accordance with rules of morality (moral rules). Hence, one who complies with the basic principles of virtue can, it is true, commit a *fault* (*peccatum*) in putting these principles into practice, by doing more or less than prudence prescribes [*MM* 6: 433].

So, if we classify a certain action as "lying", we have to prove the corresponding *Maxim* relating to this concrete action. Seen in this light, there "is" no "lying" independent of the *Maxim* of lying which is realised by this action under the concrete circumstances. For this reason it seems absurd to refer to a "lie" as "white", "excusable" or "reprehensible". As we cannot base our definition of lying on an action, but need to examine the corresponding *Maxim* by way of reflection, an action only then *is* lying, if the corresponding *Maxim* underlies. Otherwise it simply *is not!* The latter is the case, if our moral reflection results in the conclusion that under these concrete circumstances no *moral* obligation to a truthful statement exists. This means: "Lying *in general*" cannot be identified as action independent of the circumstances and *Maxims*, nor is it possible in turn to derive from the moral call for truthfulness *which* concrete actions realise it *without* taking the concrete circumstances into account.

II.3. Special cases of moral conflicts: situations of dilemma

A different thinking is required by situations in which the person concerned is actually confronted with a *moral* dilemma. A definition of such dilemma, however, also needs to take into account that the individual in the relevant situation is unable to act up to their moral convictions, due to the circumstances given. Moreover, every possible option to act comes along with tragical consequences. In this connection reference is frequently made to a scene in the film "Sophie's Choice"[26]: A mother upon arrival in Auschwitz is faced with the choice of saving one of her two children. She is forced to make a decision, otherwise both children will be killed. From the perspective of Kant's ethics, there cannot be any morally satisfactory solution of this case. In fact, a moral judgment of the person acting – whatever he or she may finally come to decide – is not even at stake. It is not Sophie who is to blame for the fact that this situation offers no option without tragic consequences, but solely those who deliberately create it or fail to prevent it. The options for acting in a situation of dilemma, as well as the actual decision made, are not subject matters of moral judgment. If anything, it is the *Maxims* of those involved in the situation, which can be morally judged. If someone intended to make use of such a situation of dilemma to get rid of an unloved child, or an inconvenient competitor, this would have to be blamed as immoral irrespective of the given parameters, which the person is not responsible for. Cases such as "Sophie's Choice" and Kant's legal example leave no doubt about the intention to pursue *Maxims* of moral integrity: Sophie wants to save both her children and, in Kant's example, friend wishes to save friend.

Against the backdrop of these thoughts the claim ascribed to Kant "you can, as you ought to!"[27] starts to make sense. We can determine our will and our *Maxims* in any situation. However, if this "can" is not taken as the capability of an autonomous determination of the will, but rather mistaken for the concrete power to realise certain ends, one extends the ethical demand and the moral responsibility to include a sphere which

26 Cf., e.g., David O. Brink, 1996, "Moral Conflict and its Structure", in: H. E. Mason (ed.), *Moral Dilemmas and Moral Theory*, Oxford: Oxford University Press, pp. 102–127, at p. 107; cf. "Sophie's Choice", directed by Alan J. Pakula, 1982.

27 Similar formulations are to be found, e.g., in: Immanuel Kant, "Die Religion innerhalb der Grenzen der bloßen Vernunft" [*R* 6: 145, 148, 150, 162].

may not at all be at the agent's disposal. The morality of a person accordingly would depend on the arbitrary circumstances and their capability to dispose, for instance, of certain material resources: technical-practical competence; or knowledge; or education. We might then be obliged to something that we cannot – in Kant's view an almost absurd and deeply inhumane idea. It is precisely this "fated" concept of commitment and obligation the Categorical Imperative is meant to free us from. Therefore, the ethical-moral quality of an action depends on the moral quality of its underlying *Maxim* and not on the agent's power to realise it. But what exactly is a *Maxim?*

III. What Is a *Maxim?*

III.1. The problem of critical self-descriptions

The term *Maxim* is often taken as the personal intention, or the motive, of an agent. According to this understanding, Kant uses the term in reference to a psychological fact, which can only be recognised by way of introspection. His ethics, therefore, seem to be of psychological-individualistic nature: a rather "private matter" that anyone can consider at their own discretion, particularly as the *Maxims* of the agent remain concealed to others, sometimes even to the agent himself. If ethical demands address the agent's intentions only, "good will" would have to be identified with the psychological will of a particular agent. The moral judgment, then, would have to assess every action as "good", if it is performed with a "good intention" and thus was "well-meant". Even if the outcome of such "good intention" itself were entirely problematic, the "good intention" alone would still ennoble the agent morally. Personal offences to the respect and dignity of others, insofar as they are not justiciable, but still are offences to private relationships, could at any time, be freely reassessed in one's own favour and thus be withdrawn from an objective moral judgement. One can always claim that one had only the best of intentions. Any act of lying could seemingly be legitimated by way of arguing that one only intended to avoid upsetting the other person. Any random help offered, even if embarrassing to the other person, would have to be acknowledged as moral action, if one's "good intention" is claimed. In setting *willing* and *acting* aside from each other, this reading of Kant's ethics, cannot make any valuable contribution to our moral guidance.

The general formulation of the Categorical Imperative, suggests a different interpretation: it does not call for one to have a *disposition* of which one can will that it became a general law, but demands: Only *act* on that *Maxim* through which you can will at the same time that it became a general law cf. [*G* 4: 421]. Kant's definition of "will", as we have seen before, determines the willing as a particular form of consciousness: one that aims at the realisation of an imagined objective. *Maxims*, moreover, should be in the nature of a principle. They should be true descriptions of the actual principles we have accepted in our acting. As principles, however, they do not relate to a person's *individual* actions or *intentions* to act, nor do they formulate concrete instructions for acting. Demanding that one should always tie one's left shoe-lace first is not a *Maxim.* Unlike the concrete intention a *Maxim* neither is given together with a certain action, nor does it have to be consciously pursued by the agent. For this reason the psychological self-description of the person cannot be decisive for the identification of a *Maxim.* It is rather our actions as objectively observable repercussions of our will that need to be taken into consideration. Accordingly, *Maxims* develop from a reflection on our acting. Such reflection, however, involves a fundamental uncertainty as to the correct assessment of both our actual *Maxims* themselves, as well as their moral quality cf. [*MM* 6: 392]. This uncertainty prompts us to seek a method which critically guides our self-descriptions and takes them towards a positivity, so that we cannot find comfort in forming moral ideologies or, as Kant calls it, an "inner lie" [*MM* 6: 429].

III.2. *Maxims* are hypotheses of the basic moral attitude of a person

An examination of the *Maxim* of acting aims at finding out what our acting *means* in a moral respect, which *spirit* is expressed by this acting. The "spirit of an action" is not a psychological phenomenon. One has to conceive it as a "moral meaning" which we can generate in reflecting on our factual or our actually intended acting. The moral meaning of an action is rendered in linking it to the social context and to other actions of the same person. In this connection one has to be aware of the fact that any acting has repercussions beyond its immediate occurrence. This is because any acting produces and fixes formations within the social setting, which could be designated as "structures". For a moral assessment of such structures, it is not the psychological internal perspective of a

person which is decisive, but solely the question of which *moral* circumstances our acting creates, or would create, for ourselves and for others.

In an attempt to capture the *Maxim* of our acting, one consequently must morally assess its formative impact on our own acting and that of others. The moral assessment examines the way in which our acting forms the "social bond" – not the psychological intentions pursued by it. If, however, *Maxims* – interpreted in this way – are gained from the reflection on our factual actions or intentions, their identification and their evaluation is not a privilege of the agent alone. This understanding lends the moral judgment an objective dimension and commits it thereto at the same time. A "disposition" manifesting itself in any single action, which consequently never materializes, is, therefore, morally wrong. For making moral judgments we have to judge realities, attitudes and basic outlooks which become manifest in actions – not potentialities and assumptions.

There is no established repertoire for "reading" the moral meaning of such manifestations, but only a method of forming such repertoire: If we try to discern *Maxims* in actions, we have to analyse and evaluate each situation from the angle of the Categorical Imperative, i.e. with respect to its possible effect on equality and respect, recognition and dignity. However, we must not lose sight of the fact that we are not faced with causal events but with actions, with acts of freedom performed by people. For this reason the moral meaning rendered by such evaluation of structures from an epistemological perspective is, by definition, preliminary in nature. It is less of an ascertainment than a hypothesis.

IV. Final Consideration: The Pragmatic Dimension of Morality

Applying this reading of *Maxims* to the problem of moral conflicts and dilemmas opens a new perspective on the present debate. It is not the concrete decision in a particular situation, not an individual action, at all, that is crucial for the morality of a person, but solely which basic principles the acting of this person reveals in its continuity. There cannot be any correct solution of moral conflicts from this perspective and, moreover, such a solution would not tell us much about the basic moral attitude of a person, anyway. Returning to the examples of the mother in "Sophie's Choice" and Kant's servant: in order to gain elucidation about the *moral meaning* of a decision someone has made, in a Kantian ethics

the hypothesis propounded with reference to this individual action would also have to hold good for preceding and subsequent actions. Thus it is possible that the judgment on the moral integrity of a person changes in view of their previous or further acting. As far as the film "Sophie's Choice" is concerned: it draws a definitely complex portrait of its main character, because it does not present her as a consistently upright victim of the circumstances, but also deals with the subject of her morally problematic decisions. So, as the story develops we have to correct our hypothesis of Sophie's *Maxims* of acting various times. What we get in the end is not a heroic character, but the realistic picture of a person absolutely struggling with the claims of morality. Still, the way the character is presented, entirely in line with a Kantian assessment of this story, does not leave any doubt about the fact that Sophie is neither responsible for the murder of her child, nor that her sense of guilt has a moral reason.

The insight we gain from Kant's ethics and his position as to the problem of moral conflicts is this: morality is neither capable of disposing of tragic situations once and for all, nor can it save the victims of such hard cases from their consequences. Equally, the cases of so-called symmetrical situations, widely discussed in current debates, cannot be reasonably solved by means of morality. (An example of such symmetrical situations would be a scenario in which twins are trapped in a burning house, but the helping person can only rescue one of them.)[28] Many questions of our acting – this is demonstrated by the Categorical Imperative – cannot be directly answered by moral duties. Often the issues do not even come within the scope of morality. According to Kant, it is, for instance, not a question of morality whether one should have a highly talented child learn to play the violin or the piano; whether one may delay the promised visit to the zoo because of an urgent appointment; or interrupt one's diet to eat a rare Indian dish.[29] Ethics, consequently, does not offer immediate advice as to how to lead one's life, or give a "general clue" for coping with any odd problem the world confronts us with. What it can do, however, is inform us in each case how

28 See Brink, 1996, p. 107.

29 See Peter Railton, 1996, "The Diversity of Moral Dilemma", in: H. E. Mason (ed.), *Moral Dilemmas and Moral Theory*, New York/Oxford: Oxford University Press, pp. 140–167, at p. 147; Elijah Millgram, 1997, "Incommensurability and Practical Reasoning", in: Ruth Chang (ed.), *Incommensurability, Incompatibility, and Practical Reason*, Cambridge/Massachusetts: Harvard University Press, pp. 151–170, at pp. 156, 159.

our acting relates to the moral demand that persons – insofar as potentially capable of self-determination – "are ends in themselves" and have a dignity. Ethics can help us to come to terms with tragic situations, because it reveals what may be *morally* assigned at all. However, the hope for establishing concrete instructions for acting in precarious situations, as harboured for instance by virtue ethics, cannot be fulfilled on the basis of moral considerations. Moral *Maxims* can be realised in various ways: "and judgment can decide what is to be done only in accordance with rules of prudence (pragmatic rules), not in accordance with rules of morality (moral rules)" [*MM* 6: note 433]. Judgment deliberates what is at the agent's disposal for realising their ethical *Maxim.* In doing so, however, it operates in ethical terms. This means that prudence considers what is morally justifiable and legally permitted for the realisation of a *Maxim.* Judgement, in this connection, is not guided by an "instrumental reason", but weighs the possibilities for realising what is morally called for. In this process, of course, the agent may commit "a *fault* (*peccatum*) in putting these principles into practice […]. But insofar as he adheres strictly to these basic principles he cannot practice a vice (*vitium*)" [*MM* 6: note 433].