Beatrix Himmelmann, Robert B. Louden (Eds.)
Why Be Moral?

Why Be Moral?

Edited by
Beatrix Himmelmann and Robert B. Louden

DE GRUYTER

An electronic version of this book is freely available, thanks to the support of libraries working with Knowledge Unlatched. KU is a collaborative initiative designed to make high quality books Open Access. More information about the initiative can be found at www.knowledgeunlatched.org

BY-NC-ND

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License. For details go to http://creativecommons.org/licenses/by-nc-nd/4.0/.

ISBN 978-3-11-037045-4
e-ISBN (PDF) 978-3-11-036639-6
e-ISBN (EPUB) 978-3-11-038633-2

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de abrufbar.

© 2015 Walter de Gruyter GmbH, Berlin/Boston
Druck und Bindung: CPI books GmbH, Leck
∞ Gedruckt auf säurefreiem Papier
Printed in Germany

www.degruyter.com

Acknowledgements

The editors would like to thank all those individuals and institutions that have supported us in one way or another.

Several of the contributions to this volume have been presented at the conference "Why Be Moral?", which was organized by the Tromsø Ethics Research Group in the fall of 2013. The group is based at the Philosophy Department of UiT, The Arctic University of Norway. Other chapters have been written for the present collection independently. We thank all the authors for their commitment to the project and the institutions mentioned for their organisational and financial assistance.

The authors and the editors of the book owe special thanks to Peter Dennis, Fellow in the Department of Philosophy, Logic, and Scientific Method at the London School of Economics, who was a copy editor as competent and reliable as one can wish for.

Beatrix Himmelmann & Robert B. Louden, May 2015

Table of Contents

Beatrix Himmelmann

Introduction

"Why be moral?" is an age-old question, which has been discussed ever since the Sophists contested the validity of moral claims. No longer, it seems, could they be considered a matter of course. Whereas the Seven Sages of Greece had stated what they deemed obvious in laconic brevity: "Nothing in excess"; "Honour the Gods"; "Be yourself"; "Yield to justice",[1] later generations felt the need of explaining and exploring what they saw as anything but self-evident. Socrates and Plato, most prominently, not only accepted the challenge but turned the tables by asserting that examination and self-examination are part and parcel of a truly human life. Accordingly, assumptions that underlie human thinking and conduct should not be taken for granted but have to be accounted for. *Logon didonai*, giving arguments in order to justify your beliefs, is considered a requirement henceforth. These endeavours to investigate alleged givens included, contra Nietzsche, right from the beginning the issue of what he called our "faith in morality".[2] A somewhat related question, "Why should I be moral?" was also high on the agenda from very early on. Plato's well-known discussion[3] still figures into today's efforts to come to terms with this delicate as well as central question. Several contributions to this collection also bear witness to the lasting influence of Plato's thought-provoking case studies and ambitious arguments.[4]

First and foremost, this collection of new essays is intended to shed light on the "Why be moral?" question from a variety of perspectives. Different approaches originating from different philosophical traditions represent the possibilities of discussing this still haunting question today. The anthology includes authors from Europe, North-America, and Israel.

It starts out by putting the meaningfulness of the question, "Why be moral?" itself to the test. In view of the fact that countless attempts to provide for a convincing answer have turned out to be futile, Dieter Birnbacher wonders whether the "Why be moral?" problem is a real problem or what Carnap and others called a "pseudo-problem". A pseudo-problem is a problem which, for some reason or another, proves to be undeserving of philosophical effort. The "Why be moral?" question, Birnbacher argues, is not a pseudo-problem in that it is trivial. Its solution cannot be considered obvious. Neither does the concept of morality cover

1 Cf. Wachsmuth/Hense 1894, p. 125 f.
2 Cf. Nietzsche 1988, Vorrede 2.
3 Cf. Plato, Rep. II, 358a ff.
4 Cf. the contributions of Dieter Birnbacher, Robert B. Louden, and Peter Schaber.

any values any individual might prioritize, nor does it imply that moral norms necessarily override all other kinds of norms. Birnbacher suggests, however, that the "Why be moral?" problem is unsolvable for reasons of principle. Either it can be answered by giving moral reasons or by giving non-moral reasons to follow morality in cases of conflict. The first option seems to be unattractive because it is question-begging. The second option is fated, Birnbacher contends, in that it also leads to a kind of circularity. Only whoever is already convinced that morality is precious and preferable will be open to non-moral justifications of the priority of the moral.

In spite of this rather sobering result, Birnbacher does not deprive those who continue working on answering the "Why be moral?" question of any hope: "The irresolvability of the problem", he assures us, "is by no means obvious". Most of the other contributors to this anthology, indeed, do suggest solutions to the problem.

Peter Schaber also aims at investigating the question, "Why be moral?" by inspecting the possible types of answers it invites. Is the "Why be moral?" question meaningless? Suspicions to that effect have been voiced, notably by H. A. Prichard.[5] Schaber argues that the question is meaningless indeed if it is understood as the question of whether we have reasons to do what is morally required. It is meaningful, though, if it is understood as the question of what reasons can justify the claim that we ought, morally, to act in certain ways. Schaber's point is that the reasons for obeying moral principles are equivalent to the reasons that give these principles their justification. Accordingly, he rejects the widely shared view that moral theories, when trying to respond to the worries of the moral sceptic, have to proceed in two steps: First, they determine what the right moral principles are; second, they answer the sceptic's question "Why should *I* follow these principles?" Instead, Schaber suggests, it takes only one step to satisfy both requirements: The justification of moral principles provides us, at the same time, with the reasons for following these principles. Though we might very well ask whether a certain moral principle is justified, we cannot sensibly ask whether we have reasons to follow a moral principle that we take to be justified.

What are the moral principles that can be justified? In what way, if any, does the kind of moral principle suggested shape the answer that is given to the "Why be moral?" question? This anthology presents a range of different responses to these queries, beginning with Robert Louden's paper.

In order to convince moral sceptics that they have sufficient reason to give up their scepticism about morality and instead try to be moral, Robert Louden

5 Cf. Prichard 1968.

considers it necessary to locate non-moral reasons that are powerful enough to do the job. He is very clear, though, about the failure of two standard candidates. Neither appeals to self-interest, nor related appeals to happiness, provide any credible reasons for being moral. The pull of self-interest itself may often be the biggest barrier to acting morally, so it would be wrong-headed from the start if we relied on one of the sources of immorality in order to solicit morality. Prospects of happiness, on the other hand, do not necessarily arise as a result of moral conduct and, therefore, cannot always motivate it. Louden's own strategy involves drawing attention to the moral norms and values that lie behind or are presupposed by ordinary cognitive activity and rational communication. There are norms, he demonstrates, regarding how we ought to think and reason, and some of these norms, he claims, are moral norms. If we can show, so the argument goes, that anyone who wishes to engage in such activities successfully must first adhere to certain underlying norms and values, then we have also shown, in effect, that anyone who wishes to engage in ordinary cognitive activity and rational communication must also try to be moral – at least in some minimal sense.

Whereas Louden and other contributors take the "Why be moral?" question to be raised by someone considering herself outside morality, Hallvard Fossheim investigates the significance the question might bear for someone who clearly is and conceives of herself as being inside morality. He does this with Aristotle as argumentative source. From an Aristotelian point of view, asking "Why be moral?" may, at first sight, look like some sort of shortcoming. It seems to be an instance of Williams's famous notion of "one thought too many" since it appears to suggest an agent alienated from or uncommitted to an ethical stance. Also from a virtue ethical standpoint, to ask this question seems to indicate some kind of deficiency. It implies that morality is not settled in you as your character. Contrary to those expectations, Fossheim aims to show that being able to ask "Why be moral?" and similar questions forms a necessary prerequisite in the process of moral improvement and for developing a moral stance that is both far-sighted and broad-minded. Hence the virtuous person's questioning and self-questioning may be seen as tokens of full ethical agency rather than moral immaturity.

Contrasting with Fossheim's investigations, Ivar Russøy Labukt explores the viability of the egoistic answer to the "Why be moral?" question. He does so, he explains, not because he finds this approach immediately attractive, but because he is convinced that it is, ultimately, the only kind of justification of morality that is available. Labukt does not, in this paper, defend his scepticism towards non-egoistic justifications. Instead, his aim is to describe and assess the egoistic alternative in order to show that it is more satisfactory than most philosophers believe. Ego-

istic considerations, he argues, in most cases support a commitment to morality that is fairly deep and at least as extensive as the one displayed by "most actual people". Labukt identifies three kinds of egoistic reasons that count in favour of a commitment to morality. Strategic reasons suggest following moral rules in order to obtain non-moral benefits or avoid negative sanctions. Psychological reasons suggest following moral rules in order to comply with what our general capacity for sympathy and the capacity for a conscience recommend doing or omitting – on pain of suffering emotional distress and missing out on substantial sources of pleasure and joy. Reasons speaking in favour of a deliberative commitment to morality suggest following moral rules in order to gain the benefits of being involved in "attractive projects", such as doing what one, on a cognitive level, takes to be morally right, or providing structure to one's life. Labukt admits, though, that there are limitations on any egoistic justification of a moral way of life: it does not apply to all persons and circumstances.

This latter point is confirmed by the results at which other contributors to this volume arrive. Responding to Hobbes, who is up to this day one of the most inspiring authors claiming that enlightened self-interest is the source as well as the motivation of being moral, Beatrix Himmelmann argues that this position is mistaken. The same holds true for Nietzsche's alternative contractualist theory, which relies on the idea of an equilibrium of power among agents. Even though both positions seem to be attractive because they do not allow for any gap separating the "real" world and "ideal" morality, they fail to do justice to the human condition. Contrary to Hobbesian assumptions, there is more to human reason and reasoning than looking out for pleasure maximization and resolving related technical problems the most important of which is taken to be the institution and maintenance of social stability. We are not merely and not essentially pleasure-maximizing or power-balancing rational animals, Himmelmann contends, pointing towards a specific use of reason that distinguishes humans: a capacity for relating to themselves by thematising themselves that they show, for instance, in referring to themselves and their doings as "I …", or by being aware of their finitude. Thus being able to become an issue for themselves, these animals must *demand* to be treated accordingly. The capacity for self-knowledge, entailing self-esteem, self-will and all the other forms of self-relation, turn human beings into very special animals who cannot claim a right to ignorance and innocence any longer. Why be moral? In order to live up to our knowledge of who we are.

David Sussman discusses the question in the light of a "constitutivist" approach. The aim is to show that commitment to moral principles is a necessary condition of being a rational agent. There would be evidence then that there is no wholly non-moral perspective any agent could hope to inhabit. This approach

can be traced back to Plato, Aristotle, perhaps Kant, and many others; more recently it has been prominently advanced by Christine Korsgaard.[6] Sussman criticizes Korgaard's account. As Sussman explains, Korsgaard argues that the formal aim of action is to integrate agents in a way that enables them to act, much as the proper functioning of any living thing serves to sustain and reproduce that organism. The self-constituting function of action involves, for Korsgaard, practical reasons that are "public" or "shareable", thus meeting the requirements of Kantian universalizability. Sussman suspects that this ideal of impersonal universality, supposedly grounding individual integrity, carries things much too far in that it does not describe fully rational agency but deprives the agent of any character that he could recognize as his own. Against this type of "alienation", Sussman sides with Bernard Williams. Williams shows convincingly, Sussman argues, that there are limits to choice and self-determination which supposedly provide assurance of remaining oneself "come what may". There are situations, however, which do not allow for any such choice but only for an acknowledgement that there is no choice left. Nobody knows or needs to know in advance where these limits lie in his or her own case. Kant's moral law could then be recovered as a regulative rather than a constitutive principle of rational agency. Asking "Why should *I* be moral?" would turn out to reveal a commitment to some principles of obligation and inter-personal accountability the scope of which will correspond to the forms of life that the agent could inhabit.

While Korsgaard's adoption of specific Kantian lines of thought is widely appreciated, it is also recognized as more or less detached from what we can find in Kant's own work. What is Kant's precise answer to the question, "Why be moral?" This is the issue Christoph Horn investigates. First, he shows why the "constitutivist" response is insufficient. According to Kant, the moral law cannot be derived from any antecedent good such as rational agency and its enabling conditions. Second, Horn argues that Kant does not accept moral intuitionism either. His idea that the consciousness of the moral law has to be seen as a "fact of reason" cannot be taken in the sense of some sort of immediate insight into moral obligation. The "Why be moral?" problem then would resolve itself. Horn's own suggestion is that Kant's answer to the question is somewhat close to the solution proposed by ancient eudaimonism, according to which moral agency is characterized as advantageous to the agent himself. By practising morality we are, on this view, actualizing our "true" (rational) nature and are gaining the happiness that is inherently connected to self-perfection. Horn's suggestion comes as a surprise given that Kant explicitly set out to untangle moral claims

6 Cf. Korsgaard 2009.

and the demands made by our striving for happiness. However, by way of analysing Kant's concept of will, which turns out to imply features of a pre-modern understanding of this faculty pointing towards qualities of purity and purification, Horn arrives at what he takes to be Kant's (indirect) answer to the "Why be moral?" question. Following the moral law ultimately leads us to the only appropriate sort of happiness: that which is qualified by our moral worth. By being forced to neglect our happiness, Horn argues, the moral law realigns us – and precisely thereby, it finally brings us to our "true" nature and happiness.

Iddo Landau chooses an interesting and related angle in that he links the question, "Why be moral?" to the notion of the meaning of life. Overall, he aims to show that morality makes our lives more meaningful. He proceeds by invalidating all kinds of objections that might be raised against this claim. Critics could point out that some theories of the meaning of life are subjectivist, and as such imply that meaningful lives need not be moral at all. But various considerations, Landau argues, suggest that subjective theories of meaningfulness are too problematic to accept. This is one of them: Given that meaningfulness rests on worth or value and meaninglessness, correspondingly, on the lack of worth or value, people can be wrong in their evaluations, including their self-evaluations. So it looks as if a more convincing account of meaningfulness need to be objectivist to a certain degree and may include, to a certain extent, features of moral commitment. Critics, however, may accept this move while denying that these objective features have to do with morality. Some theories of the meaning of life present both subjective and objective conditions for meaningfulness and do not mention morality at all, thus allowing for highly immoral lives to be considered meaningful. But if a meaningful life, Landau insists, is a life that, overall, has a sufficiently high degree of worth or value, it cannot be very low in, say, morality, yet very high in meaningfulness. So morality and meaningfulness are not independent of each other. Morality affects the overall value of one's life and thus its meaningfulness. Also, since the notion of a meaningful life is a laudatory, honorific notion that has positive connotations it cannot be applied to individuals committing horrific crimes and saying they find meaning in so doing, whereas we would consider their immoral behaviour despicable or abhorrent. Landau acknowledges, though, that lives can be highly meaningful without displaying moral excellence. Some great artists may serve as examples. So it seems as if refraining from behaving in highly immoral ways is a necessary condition for having a meaningful life, whereas behaving in highly positive moral ways is not. Why be moral? Because it frequently enhances meaningfulness, Landau replies.

This is an answer which Roe Fremstedal, discussing Kierkegaard's approach, would rather make stronger: we should be moral because it is essential to mean-

ingfulness, full stop. Not least differences in identifying the demands of morality may lie at the heart of differences in specifying the demands of being moral. On the one hand, there are conceptions distinguishing some minimal sense of morality (and being moral) from richer or thicker notions of morality (and, correspondingly, higher levels of the moral commitment required), thus allowing for quite a range of different degrees of morality (and senses of being moral). On the other hand, there are conceptions which do insist that morality, by its nature, does not allow for such grading, thus ruling out the idea of minimal or maximum senses of morality (and being moral). This same division holds true for most of the contributions to this volume.

Fremstedal presents Kierkegaard's multifaceted ideas, as well as contemporary versions of them, by bringing out the main arguments they involve. He principally, but not exclusively, focuses on the debate between various aesthetes and the ethicist as depicted in *Either/Or,* concentrating his analysis on the position of the reflective aesthete. Themes of central significance are love, selfhood, and freedom. The ethicist argues that it is in the aesthete's true interest to become an ethicist. Without ethical commitment, he contends, love is episodic, lacking continuity and importance, selfhood unbalanced, and freedom negative, empty, and arbitrary. Thus, the ethicist confronts the aesthete with an *internal critique* rather than condemning the aesthete on moral grounds, something that may be seen as moralistic and unhelpful. The aesthetic stage seems to fail on its own terms, and it appears to be preserved in the ethical stage. To defend these claims, Kierkegaard chooses to develop a *via negativa:* we only understand the ethical through its failure, through guilt, sin, and despair. Fremstedal elaborates on "the argument from despair" in detail and concludes by highlighting the role of hope and honest commitment. In his *Works of Love* Kierkegaard argues that hope without the moral duty to love one's neighbour is false. Hoping only for myself involves conceiving of hope as something private that does not concern my relationship to others, as if I could have a future of my own without others. Hoping in this way fails to appreciate the extent to which I am dependent on others. Also, avoiding honest commitment, keeping things and people at a distance deprives the aesthete, finally, of any stance, identity or character whatsoever. He would be doomed to despair.

Richard Eldridge suggests another, modest way of justifying morality and being moral. It is not through despair that we come to understand and acknowledge what Eldridge – loosely following Kant – calls the moral law. Starting out from Bernard Williams's questioning of the absolute authority of distinctively moral reasons, Eldridge aims at elucidating how moral reasons make claims on us. This task, Eldridge points out, requires accepting the fact that we are embedded in diverse personal and social relations. The emergence of distinctively

moral commitments out of an enormous variety of developmental backgrounds, both individual and social-historical, may display a kind of path independence. Normative reasons are instituted by us in the course of our complex practical lives, and they are subject to historical variation. What is deemed morally right or wrong is often a matter of dispute. For instance, what one of us calls telling someone a hard truth out of respect another one calls inconsiderate cruelty; what I call encouragement to develop one's talents and specific forms of self-respect you call indulgence and pampering. Mistakes and errors on all sides are possible. What counts as respect for persons is itself a subject for open, imaginative, explorative inquiry and moral conversation in an ongoing way, even where the value of respect for persons is abstractly shared. What reasons, finally, do we have to acknowledge the moral law? We cannot have, Eldridge concludes, but happily do not need, a justification for being moral "from nowhere", apart from our location within a set of developing circumstances of life.

This is a subject-matter which is also touched upon by Erik Lundestad. His essay enquires into the argumentative resources that pragmatism may contribute to the discussion of the "Why be moral?" question. Pragmatists hold, as Lundestad shows, that the question both can and ought to be dismissed. As they see it, this question only appears interesting to those of us who, mistakenly, assume that morality forms a sphere distinct from that of prudence. Given this misconception, moral acts will be seen as having a specific aim, different from that of prudential acts. Pragmatists, however, dismiss the idea that there is any such gap between morality and prudence. This distinction, they argue, does not have any precursor in practice, in the manner in which we act. Everything we do is aimed at a good; the question, "Why be moral?", then, does not seem to be very exciting. Of course, pragmatists have to reconstruct or redescribe what is usually perceived as a specifically moral act if they want to be convincing. Focusing his analysis on the work of John Dewey, Lundestad presents the main arguments of pragmatism and confronts them with pieces of criticism put forward by Habermas and others. One of the most serious problems arising is the risk of conflating what *is good* with that which a contingent society at a given time actually *takes to be good*. While Dewey approaches morality from the perspective of community members concerned with their common good, Habermas contends that this perspective is mistaken. It is only on the basis of a distinctive "*horizontal* We-perspective" and not a "*vertical* We-perspective" that we are able to account for morality. Thus it may well be true, Lundestad concludes, that it isn't the "Why be moral?" question but pragmatism's dismissal of it that has to be dismissed.

Héctor Wittwer discusses another option of setting aside the problem. This option would be available if it could be shown that the problem is trivial because

its solution is obvious. "Is the overridingness of moral reasons a semantic fact?", Wittwer asks, and he carefully analyses different aspects of this question. If it could be given a positive answer, it would be clear from the outset why each of us should be moral: because moral reasons for action always and systematically override all other kinds of reasons for action. At least, this understanding would be implied by any speaker who properly uses terms such as "moral" and "morality". The (normative) overridingness of moral reasons would be simply a part of the language-game of morality and, hence, a *semantic fact*. Wittwer refutes this claim, presenting a variety of objections. There is no consensus about the precise meaning of the word "morality" among all competent speakers of the different languages, and it cannot be established either. We would not necessarily consider sentences such as the following contradictory: "Even though it is your moral duty to do *x*, you should refrain from so doing because it is imprudent." For these (and still other) reasons, normative overridingness cannot be a conceptual feature of morality, Wittwer concludes. But we can neither provide any non-linguistic justification of the alleged overridingness of morality, he contends. Should the claim pertain to the relative priority of morality, as a consequence of assessing things from the moral point of view, it is true but trivial. If it refers to the absolute priority of morality, the argument is doomed to fail, Wittwer assumes, because "nothing can ever have absolute priority over something else". So it looks to him as if, finally, we are left with nothing but the (rational) preferences of agents, choosing between self-interest and moral demands. Whatever they decide to do, no normative conclusion, he thinks, can be drawn from their actual choice.

Do we really have to leave the problem behind like this? Even though Bernard Williams, as mentioned before, doubts the absolute authority of moral reasons and disapproves of what he calls the "morality system",[7] he does offer an – ambitious – answer to the question, "Why be moral?" This is shown by Alan Thomas, who addresses the question from the standpoint of Williams's moral and political philosophy. He extends his consideration of Williams's answer from its origin in a conception of practical reasons – as always "internal" – to Williams's political psychology, particularly focusing on the late discussion of this topic in *Truth and Truthfulness*.[8]

Certainly, Williams suggests no formal answer involving a priori constraints on practical agency as such. As Thomas expounds, the question, for Williams, can only have a substantive answer; that is it can only be answered in the

7 Cf. Williams 1986, Chap. 10.

8 Cf. the bibliography.

light of the substantive content of an agent's "ground projects". Hence, those answers fall within the class of "moderate morality"[9]. Furthermore, Thomas argues, it can be shown that Williams considers morality as consonant with a background conception of ethical community that involves relations of mutual respect. This is the upshot of the argument Williams pursues in his late book *Truth and Truthfulness*. In complex and interesting ways illuminated by Williams, "Accuracy" and "Sincerity", both epistemic virtues, are shown to have shaped the "system" of knowledge, disposing people *to take care* and *not to lie* or otherwise mislead. This "politics of knowledge", as Thomas demonstrates, helps to explain the phenomenon of a "politics of recognition", its ethical (and political) complement. While individuals develop identities by committing themselves to "ground projects", which are as much found as made, they call for *acknowledgement* when presenting themselves to others. An essential role is played by mutual recognition and the very sociability that makes individuality possible. Any answer to the question, "Why be moral?", Thomas concludes, is constrained by mutual acknowledgement and mutual respect in the context of a modern ethical community, which has to be seen as a historical achievement.

Presumably, none of the contributors to this volume would disagree.

Bibliography

Korsgaard, Christine M. (2009): *Self-Constitution: Agency, Identity and Integrity*. Oxford: Oxford University Press.

Nietzsche, Friedrich (1988): "Morgenröte". In: Giorgio Colli/Mazzino Montinari, eds.: *Kritische Studienausgabe* (= KSA 3). Berlin/New York: De Gruyter.

Prichard, H. A. (1968): "Does Moral Philosophy Rest on a Mistake?" In: *Duty and Interest*. Oxford: Oxford University Press, pp. 1–17.

Scheffler, Samuel (1992): *Human Morality*. Oxford: Oxford University Press.

Wachsmuth, Kurt/Hense, Otto, eds. (1894): *Ioannis Stobaei Anthologium,* vol. 3. Berlin: Weidmann.

Williams, Bernard (1986): *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Williams, Bernard (2002): *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.

9 Cf. Scheffler 1992, p. 6.

Part I: **Investigating the Question**

Dieter Birnbacher

Why Be Moral – A Pseudo-Problem?

1 Introduction

The question what reasons we have to do what is morally required instead of what we happen to want for nonmoral reasons is one of the oldest questions of moral philosophy. It looms large already in the Platonic dialogues, especially the *Gorgias* and the *Republic*. At the same time it is one of the most controversial problems. This is true not only in the sense that there is controversy about which of the solutions proposed for it has the best chances to be true or adequate but also in the sense that it is an open question whether it is a real problem or what Carnap and others called a "pseudo-problem", i. e. an apparent problem which, for some reason or other, proves to be undeserving of philosophical effort. Though the Why-be-moral-problem is only rarely given the label "pseudo-problem" (but see Thornton 1970, p. 453), there is a presumption that it is in fact a pseudo-problem given the futility of the numerous attempts to solve it.

A problem may be classified as a pseudo-problem for a number of reasons: that standard formulations of the problem are semantically *meaningless* or otherwise *unintelligible*; that statements of the problem start from *wrong* or *contradictory presuppositions*; that it is *trivial* because its solution is obvious; or that, to the best of our knowledge, it is *unsolvable* for reasons of principle, so that further attempts at a solution can be expected to be futile. Not all of these justifications for reducing a problem to the status of a pseudo-problem are relevant to any purported pseudo-problem. A paradigm case, however, for which they are all relevant is the problem to which Heidegger gave the name "fundamental problem of metaphysics", i. e. Leibniz's question why there is something rather than nothing. There are good grounds to think that the *Grundfrage* is in fact a pseudo-problem though not for the reasons given by Wittgenstein and the Logical Empiricists (cf. Birnbacher 1990). In many ways, the role the Why-be-moral problem plays for moral philosophy is comparable to the role the "riddle of existence" (Rescher 1984) plays for metaphysics. It is a foundational question which nevertheless is only rarely made the focus of philosophical inquiry, not least because of its controversial status.

The aim of the following contribution is to see how far the suspicion that the Why-be-moral problem is a pseudo-problem can be supported by applying strategies and arguments similar to those that have been deployed in throwing the *Grundfrage* from its metaphysical throne.

2 The Triviality Thesis

One strategy to expose the Why-be-moral problem as a pseudo-problem is to argue that it is trivial because the solution of it is analytically implied by the terms by which it is formulated. The term that is supposed to do the job is the term "moral". This term, according to the triviality thesis, prejudges the question why morality should have priority and thus renders any further reflection about the Why-be-moral problem superfluous.

There are, as far as I can see, two variants of this thesis. The first variant was put forward by the Australian philosopher D. H. Monro. According to Monro, there is no reason to ask why we should do what is required by morality because the concept of morality covers anything that we prioritize in our personal aims. "Morality" refers to the values that are most important to us, no matter whether these are egoistic, altruistic, or impersonal, whether these values are universal or idiosyncratic, constant or transitory. Morality, thus understood, stands for an individual's highest priorities, irrespective of their content, their temporal stability and the degree of the individual's commitment to them.

At first sight, Monro's position looks like a sleight-of-hand. It is, however, not as extreme as it may appear. It does not imply that we necessarily follow morality because whatever we do is directed at the realization of our highest values. Instead, it reserves the name "morality" for whatever we *think* are our highest and most important values. Thus, it leaves room for *akrasia* and other failures to conform to these values in our actual behaviour. Morality is not the name for what in fact drives us in our actions but what we think *should* drive us. In this way, the question, "Why should I be moral?" is not exposed to the criticism of meaninglessness. There may be a considerable gap between a person's priorities (whatever they are) and a person's factual choices. Nevertheless, it may be doubted whether admitting so much reduces the triviality of the question. It is evident that a person always has a good reason to do what the person thinks it is important to do, and that the best reason a person can possibly have for a course of action is that it is the course that is, given the background of this person's values, the most important. Therefore the question why I should do what I think I should do has a trivial answer: Nothing could be more obvious.

This result, however, essentially depends on the acceptability of Monro's explication of the concept of morality. Whence comes this concept of morality? Monro appeals to common usage:

> We sometimes use 'morality' of any over-riding principles, whatever their content ('Satan's morality, 'his morality is purely selfish'). (Monro 1967, p. 225)

In a similar vein, Kurt Bayertz thinks that this comprehensive use of the term "morality" is at least "one of the ways the word 'morality' is used in everyday language":

> In a first use, "morality" designates a complex of norms, values or ideals that provides each individual with a general orientation for leading his or her life. (Bayertz 2004, p. 34)

Is this plausible? I think the answer has to be negative. Were we to accept Monro's and Bayertz's thesis, we should be compelled to give the name of morality to absolutely whatever an individual puts in the first place. This would not only be incompatible with widespread usage, it would be positively misleading. Not only would the amoralist whose maxims glorify pure egoism or arbitrariness have to be credited with a morality. The same would have to be done even in the case of the immoralist who deliberately breaks the most essential moral rules (such as the character Lafcadio in André Gide's *The Caves of the Vatican*). We have to accept, I think, that morality is not to be equated with any system of norms, values and ideals a person happens to adopt or to follow. It is something distinctive with specific formal and material features. Of course, it is no easy task to define these features in a general and coherent way. Most of the properties attributed to morality as necessary and/or sufficient conditions do not stand the test of confrontation with counterexamples. On the contrary, the failure of the search for strictly defining conditions of the concept of morality (see, for example, Wallace/Walker 1970) suggests that the characteristic features distinguishing moral from other (aesthetic, personal, or technical) norms and values are to be conceived not as necessary and sufficient conditions but as *family resemblances* in the Wittgensteinian sense or as conditions characterizing *typical* instances of the concept, i.e. as features that do not claim strict universality.

In a recent publication, I proposed four interrelated features that might be candidates for typical features of moral norms and values in this sense: *community-relatedness, obligatoriness, subjective importance*, and association with *emotions* (Birnbacher 2013). It is easy to show that individually, these four features are neither necessary nor sufficient: Not all moral values are community-related, and many sorts of norms are community-related without being moral norms. A morality can entirely consist of ideals or virtues that imply evaluations of behaviour, motives and character traits without containing or directly implying norms or obligations. Neither is obligatoriness by itself distinctive of morality. Again, moral demands are typically, but not necessarily, associated with an emphasis and seriousness that is foreign to prudential, aesthetic and conventional demands. This feature explains, at least to a certain extent, the inherent tendency of moral principles to objectification (Mackie 1977, p. 30 ff., Blackburn 1986, p.

181 ff.), the tendency to be projected into entities beyond human subjectivity and interests, into the "nature of things", transcendent values, or God's will. A further typical feature is the close association of moral norms and values with moral emotions, each of them having an important role in motivating moral behaviour.

Even if the above list of the typical features of morality is only provisional, it suffices to make it probable that Monro's thesis is wrong and that morality is not to be equated with any system of norms and values whatsoever by which a person orients his or her life. Even if morality is one, and a very important, system of norms and values by which we orient our lives, it is not the only one. Morality is a *distinctive* system of norms even if it lacks a clear-cut demarcation against other systems of norms, whether individual or social. It follows that the question why we should follow *these* norms and not others cannot be trivial.

According to the second variant of the triviality thesis the answer to the question, "Why be moral?" is trivial not because of the indistinguishability of morality and other kinds of norms but because of another assumed defining feature of morality, the *overridingness* of moral norms in contrast to other norms. The thesis is a natural consequence of the view that "overridingness" is an inherent feature of morality and that, therefore, it is analytic that moral norms claim priority in situations in which a person is in doubt whether to follow the moral course rather than a course suggested by egoism, social convention, or spontaneous feeling. The thesis of the overridingness of morality has a respectable provenance that reaches back to the Platonic Socrates. There is a suggestion of overridingness also in the Kantian tendency to associate morality with *Unbedingtheit*, unconditionality. The most explicit statement of this position is to be found in the writings of the British philosopher Richard Hare:

> There is a sense of the word "moral" (perhaps the most important one) in which it is characteristic of moral principles that they cannot be overridden [...], but only altered or qualified to admit of some exception. This characteristic of theirs is connected with the fact that moral principles are [...] superior to or more authoritative than any other kind of principle. A man's moral principles, in this sense, are those which, in the end, he accepts to guide his life by, even if this involves breaches of subordinate principles such as those of aesthetics or etiquette. (Hare 1963, p. 168 f.)

This variant of the triviality view differs from the first one in maintaining that overridingness is a necessary and sufficient feature of morality without at the same time making morality a purely formal concept. For Hare, unlike Monro, there can be nothing like a "Satan's morality". Morality has material features beyond its formal ones. Nevertheless this variant shares with the first variant the view that moral norms and values are necessarily of supreme importance and

prejudge the issue of which of two or more conflicting kinds of norms is given priority: Moral norms are by their very nature the most important norms and claim a privileged position so that the question which kind of norms should be given priority in cases of conflict does not arise.

This variant is exposed to two sorts of criticisms, each of which, I think, makes it lose much of its attraction. The first is that the inherent overridingness of moral norms and values is far from evident. Though moral norms are in most cases important considerations to be taken account of in one's choices it seems mistaken to say that they are by necessity the most important norms in any situation. There are situations in which nonconformity with moral demands, whether social or individual, seems excusable so that conflicts between the moral course of action and the prudential course of action cannot be categorically ruled out. In these situations, at least, questions of the sort "Why follow the moral course instead, say, the prudential one?" are anything but trivial. Amartya Sen gives the following example:

> There is nothing particularly schizophrenic in saying: "I wish I had a vegetarian's tastes, for I disapprove of the killing of animals, but I find vegetarian food so revolting that I can't bear to eat it, so I do eat meat." (Sen 1974, p. 63)

It does not at all seem obvious that moral principles are, as Hare claims, "superior to or more authoritative than any other kind of principle," or that moral values necessarily take precedence over nonmoral values. Nor is it obvious that "overridingness" is an inherent feature and part of the very *meaning* of morality. It is true, most people in fact see themselves as moral evaluators. But they do not see themselves as moral evaluators throughout. Their evaluations of their own and other's behaviour take many forms, and these vary with context. In many contexts – e. g. the economic one – they orient themselves primarily by prudential norms of individual rationality, in others – e. g. the family – by "self-referential altruism" (Mackie 1977, p. 84), i. e. with partiality for the near and dear. People usually live in a number of different normative worlds, a moral, a social, a prudential and an aesthetic world. Of course, from the perspective of the "moral point of view", we necessarily give priority to the moral aspects of a given case. But from this it does not follow that we necessarily give priority to this point of view. Even if moral norms and values are associated, as a rule, with particularly strong and sustained emotions, such as guilt, shame and indignation, this does not imply that whoever accepts these values gives them priority. The selfsame prominent role that moral emotions like guilt, shame and indignation play in the decidedly moral person may be played by emotions like envy, anger and pride in the rational egoist.

The second criticism of this variant of the triviality thesis is that even if overridingness were a necessary feature of morality this would not render superfluous or trivial the question whether it is followed, *in toto* or in an individual case. Normative priority does not entail factual priority. Even the most authoritative and absolute moral norm cannot prevent its adherents from acting against it. There is no pre-established harmony between the norms and values held by a person and the person's actions. It is perfectly thinkable that someone sincerely holds a certain moral principle and frequently acts against it. Espousal of the principle is not necessarily invalidated by deviant behaviour. That there is a considerable "gap" between judgment and performance is also supported by moral psychology. Empirical evidence strongly suggests that the capacity to make moral judgements is largely independent of the readiness to act in accordance with them (cf. Montada 1993, p. 268). Though internalists about moral motivation will dispute the possibility of accepting a moral rule without integrating it into one's moral outlook, this leaves open the possibility that the internalization of moral norms and values is too weak to counteract motivationally stronger nonmoral impulses and desires. I think that internalism is perfectly right in maintaining that the adoption of moral norms and values cannot be seen as a purely cognitive affair but has emotional and motivational components that dispose whoever adopts these principles to act in accordance with them. But though a firmly internalized moral conviction will, as a rule, be accompanied by a certain conformity in action, this does not exclude that it is overruled, in certain situations, by nonmoral motivations.

The conclusion of this section is negative. The examination of the two variants of the triviality thesis has shown that the hypothesis that the Why-be-moral problem is a pseudo-problem cannot expect support from this side.

3 Moral Reasons for Morality?

Are there other avenues to the conclusion that the Why-be-moral problem is a pseudo-problem? One candidate is the thesis that the problem is unsolvable for reasons of principle and that, therefore, we should discontinue attempts at a solution and rather leave it as it is.

The thesis that the problem is unsolvable usually goes together with the construction of a dilemma, both horns of which are said to lead to an impasse. The dilemma can be stated as follows: Either the Why-be-moral problem can be answered by giving moral reasons or by giving nonmoral reasons to follow morality in cases of conflict. According to the futility position both horns are equally unattractive, though for different reasons. The first horn is unattractive because it is

ultimately question-begging: the question why we should follow the moral option is reiterated at a higher level. The second horn is fated by the inability of purely nonmoral motivations to support even a minimalistic morality deserving the name.

The first horn of the dilemma is connected with some of the most well-known attempts in the history of philosophy to give morality a privileged normative status, especially over against prudential ones. Moral philosophers with the most diverse background theories have concurred in postulating that in cases of conflict between morality and prudence, morality should be paramount. If prudence or other ego-centered principles are legitimate, they are legitimate only as instrumental principles or because they are otherwise sanctioned by morality, e. g. as a rights (such as liberty rights or rights to privacy), or as mandatory rights that bind the right to pursue one's own ends (such as those directed at the realization of one's potentials or at self-perfection) to self-directed moral obligations

The problem, however, is that the principles by which this superiority of morality over prudence and other norms and values is argued are, as a rule, themselves moral principles, though of a higher generality than the principles of the morality in question. One such principle is the principle of *universalization*. It implies a strict reciprocity of the morality we expect from others in their dealings with us (and those we care about) and our moral obligations to these others. A familiar example is truthfulness. As far as I expect others to follow the moral rule to be truthful to me and people I care about, I should be prepared to follow the same rule in my dealings with others even in cases in which I have a strong interest in not telling the truth. The problem is, of course, that universalization in this sense is itself (or depends on) a rather strong moral principle, the principle that moral reasons hold irrespective of the identity of persons and that statements of these reasons must not involve proper names and indexicals. In this way what at first sight looks like a merely logical requirement develops a considerable normative force – comparable to that of moral principles like equity or fairness that are less likely to obscure their genuinely moral character. In fact, the genuinely moral character of the principle of universalization is sometimes concealed by confusing it with a principle that is in fact truly logical, the principle of the supervenience of moral properties over nonmoral properties. This principle says that situations identical in all morally relevant nonmoral characteristics must be identically judged. Given that certain nonmoral characteristics of a situation are morally relevant, they must be morally relevant in all cases and not only in a particular case at hand. But this principle is a mere principle of consistency. A breach of it would involve one in a contradiction. The principle of universalization goes beyond such a mere consistency principle in restricting the range of characteristics that are allowed to count as rele-

vant. It requires that only those characteristics count as morally relevant that can be expressed by logically universal terms. Differently from the purely formal principle of consistency, this principle is a substantive moral principle.

This point can be generalized. If the meta-principles by which an ethical theory attempts to justify the priority of moral over against other principles are themselves moral principles, this may be an important undertaking, for example to show, in formal terms, the unity and coherence of a certain system of morality, or, in material terms, the justice, fairness or efficiency with which it distributes rights and duties, advantages and disadvantages, benefits and costs over its addressees. It is unsuited, however, to give a satisfactory answer to the Why-be-moral problem. The answer it gives is necessarily incomplete because it depends on the prior acceptance of the meta-principle invoked. If neither the amoralist (who is indifferent to morality as a whole) nor the immoralist (who is opposed to it) has a reason to accept and to act on a given morality, they do not have a reason either to accept and to act on the meta-principles by which these are legitimized.

It is surprising to find that, among the moral philosophers conscious of the difficulties posed by the first horn of the dilemma, one of the most clear-sighted was Kant, the same philosopher who was more eager than anyone to establish the supremacy of moral norms and values over norms and values of other kinds. Kant saw clearly that proving the Categorical Imperative to be the fundamental principle of morality is not sufficient to establish this Imperative as the highest and most authoritative principle of action. If the Categorical Imperative is to be more than the basic principle in a reconstruction or explication of what it means to judge and act morally, more must be shown, namely that this principle is supreme and deserves to take precedence, in our judgments and actions, over prudential or aesthetic principles. If we are free, as Kant emphatically asserts, the question is inevitable why we should, in a case of conflict, opt for following the Categorical Imperative rather than for following our best interest or our strongest desire. Why should all reasons that are not strictly moral reasons be contrary to practical reason? Or rather: Why should practical reason be so narrowly defined as to admit only of one interpretation, that of conformity with the Categorical Imperative? That Kant was aware of the necessity to say something more about the authoritativeness of the Categorical Imperative than is implied by its simple statement is documented by his many attempts to provide arguments that might close the gap. One of these arguments is the argument, in his *Lecture on Ethics*, that freedom (which, for Kant, implies conformity with practical reason) is the "essential end of humanity":

> He who subjects his person to his inclinations, acts contrary to the essential end of humanity; for as a free being he must not be subjected to inclinations, but ought to determine them in the exercise of his freedom; and being a free agent he must have a rule, which is the essential end of humanity. (Kant 1963, p. 122)

It is evident that this argument is vulnerable. For one, it is not clear what "essential end" can mean in this particular context. Understood in a literal sense, an end is necessarily an end of a subject. But who is this subject of this end? In traditional metaphysics, this subject was supposed to be God or Nature. But this view can hardly be attributed to Kant, particularly in view of the personalization involved in either of them. "End" therefore should rather be understood as something like "inherent teleology". Kant seems to think that man is destined to be free, not in the sense of being literally created for this end, but in the sense of being free as part of his peculiar essence. Nevertheless, the difficulty arises that it is left unexplained why the individual should be obliged to follow this teleology. After all, non-conformity with this essence might be taken to be an even more convincing proof of his freedom than conformity. However that may be, it is clear that the kind of argument by which Kant attempts to confer legitimacy on the supremacy of freedom (*alias* morality) is itself a moral argument, an argument based on an essentialist morality.

Another argument designed to show the supremacy of moral over other kinds of principles can be found in the *Grundlegung:*

> ... because the world of understanding contains the ground of the world of sense, hence also of its laws, hence is immediately legislative in regard to my will (which belongs wholly to the world of understanding), and hence must also be thought of wholly as such, therefore as intelligence I will cognize myself, though on the other side as a being belonging to the world of sense, as nevertheless subject to the laws of the first, i.e., to reason, which in the idea of freedom contains the law of the understanding's world, and thus to autonomy of the will; consequently I must regard the laws of the world of understanding for myself as imperatives and the actions that accord with this principle as duties. (Kant 2002, p. 70)

According to this argument, practical reason deserves priority because it is part of the intelligible world while desires are part of the sensible world, with the intelligible world dictating the laws of the sensible world and therefore rightly claiming priority. Of course, the premise that the intelligible world dictates the laws to the sensible world is closely bound up with the Kantian scheme of the priority of the noumenal world in relation to the sensible world, and therefore with his own two-world metaphysics which by itself cannot claim authority. But even if one accepts this piece of Kantian metaphysics it is not evident that what is metaphysically prior is also normatively prior. Morality is a phenomenon of the sensible world, the world of human society. It might be argued that its

foundations should be sought in the facts of the world in which it comes into play, i.e. the social world, and not in the depths of a metaphysical theory. But again, however that may be, it is evident that this argument, like the former, once more legitimizes the supremacy of morality by a moral argument.

At this point, a concession is in order. The difficulties just noted in justifying the priority of morality by moral reasons apply only if the question, "Why be moral?" is understood as a question relating to moral *action* and not to morality as a *character trait*, a complex of *virtues* or as moral *motivation* in the sense of Kant's "good will". If morality is understood, not as a system of rules but as a personal disposition to act morally, a straightforward justification can be given for being moral, namely in terms of a given morality, understood as a system of moral principles. Such a justification neither begs the question nor ends up in circularity. It simply refers to the instrumentality of the possession and acquisition of moral dispositions for moral action. Given that one accepts a certain system of moral rules and is willing to act in accordance with them, it is evidently advisable to acquire fixed dispositions to follow them, both in daily life and in existentially crucial situations. Who is virtuous by habit (has made virtue his "second nature") will be spared the effort to bring himself to do what virtue demands in any concrete situation. For him, in fact, the question, "Why *be* moral?" – in contrast to the question, "Why *do* the moral thing?" – has a clear and philosophically unproblematic answer: because it disposes him to do the moral thing more easily and naturally than in the contrary case. Of course, this kind of legitimation of being moral will impress neither the amoralist nor the immoralist. It presupposes what they call in question, the bindingness of morality as a system of rules for action.

4 Nonmoral Reasons for Morality

A dilemma is complete only if not only one but both horns lead to an impasse – if there is no satisfactory answer to be expected for either option. The answer to the first horn was negative: If the question, "Why be moral?" is given a moral answer the question is only renewed. Given that the amoralist and the immoralist have no reason to be persuaded by considerations of morality they will be persuaded by considerations of second-order morality as little as by considerations of first-order morality. Are there chances for a satisfactory solution based on nonmoral rather than moral reasons? Can the egoist, the amoralist or the immoralist be persuaded to do the moral thing from prudence and other purely selfish motives?

The literature on the reasons the egoist has to act morally, from Plato through Seneca to Kant and modern moralists, is not only vast but has, as Nietzsche would have expressed it, an unmistakable smell of moralism. Though only few have gone to the Stoic extreme to assert that morality and well-being coincide, so that a loss of morality is by the same token a loss of well-being, this literature persistently exaggerates the sacrifices an amoralistic life strategy imposes on the egoist. Thus, Kai Nielsen, in one of the by now classical articles on the topic, writes:

> Given the world as it is, a deliberate, persistent though cunning policy of selfishness is very likely to bring on guilt feelings, punishment, estrangement, contempt, ostracism, and the like. A clever man might avoid one or another of these consequences but it would be very unlikely that he could avoid them all or even most of them. (Nielsen 1970, p. 477 f.)

No doubt, Nielsen has a point. There are sacrifices egoists, amoralists and immoralists have to make if they want to pursue their life strategies consistently. At the same time, one wonders how Nielsen can be as sure as he is that these sacrifices cannot be more than compensated, from the perspective of these characters, by what they gain from their immunity to moral demands.

Nielsen combines, in the above quotation, two very different kinds of "sacrifices" the amoralist has to face: the sacrifices in well-being by non-conformity with one's own personal moral convictions, and the sacrifices in well-being by non-conformity with the morality of one's various reference groups or society at large. All these moralities may differ substantially, both in content and in demandingness, and the question arises in relation to which of them the amoralist has to face the necessity of "sacrifices" of the sort referred to.

One thing that can be said at the outset is that there will be little occasion, for the amoralist, for "sacrifices" *in foro interno* from his or her own moral principles. Why should the amoralist have guilt-feelings if his own conscience is silent on his own behaviour, including what moralists would call wrong-doing and recklessness? All the amoralists cares about is how far his behaviour may be imprudent, foolish or self-destructive, and it is well imaginable that he mentally suffers from his occasional or frequent failures to act according to his best interest more than the moral person from his or her own moral lapses. But even if the amoralist suffers from pangs of (prudential) conscience for irrationality, this is far away from the guilt-feelings ascribed to him by Nielsen and others.

What about the "sacrifices" in social contacts and the lack of social respect and support the amoralist has to confront? Do they provide the amoralist with reasons sufficiently strong to make him change his ways and to follow morality in cases where this conflicts with his interests, wishes or feelings? Do the amor-

alist's motives frequently, or even necessarily, lead to estrangement from his or her social environment? This is certainly to be expected in a social environment of resolved moralists, for example if the amoralist marries into a family with a high level of moral commitment. In order to arrange him- or herself with this environment the amoralist will often have to simulate emotions and motives he does not have and dissimulate emotions and motives he has. In general, however, the psychological burden the amoralist has to bear by these manoeuvres will be limited to the extent that he manages to stay away from social milieus with highly developed moral sensibilities and a dense network of social control. Furthermore, the more open and liberal the society is in which he lives, the less likely he will be expected to show particular moral fervour. To find acceptance, he will need hardly more than to endorse, in words and deeds, the central values of his society and observe the norms of political correctness. He will be well advised to seek his contacts primarily in circles of amoralists like himself, and the more open the society is in which he happens to live, the less difficult it will be for him to find them.

The lives of amoralists (egoists included) might be a great deal better than the moralist is prepared to recognise/ize. Only the immoralist with his ambition to deliberately act against (public) morality will risk significant punishment and ostracism, at least to the extent to which he deliberately lays open his doings to public inspection. But the immoralist in this sense is mainly a literary figure, or a construction from Nietzschean and Gidean fantasies. In the real world, amoralists and egoists are much more numerous, and they will, if only they are moderately rational, guard against public censure and the social isolation following from it. Few, for example, will risk being publicly sued for failure to pay their taxes and to suffer the loss of reputation going with it. In former times, in which the bank secret was regarded as something like the first commandment of international commercial relations, this was decidedly different. In circles of well-to-do citizens, tax evasion used to be, until recently, not only tolerated, but respected and recommended as a more or less obvious duty of self-care.

Moralists referring to the "sacrifices" incurred by amoralism often overlook a crucial dimension of the amoralist's life-balance: the compensation of potentially lacking moral respect and support by other forms of social respect and support. In many cases, the amoral person is able to gain respect and support in virtue not of moral virtue but of other qualities, such as intelligence, creativity, courage and in virtue of economic, scientific, artistic or political achievement. Often, these qualities are further garnished by personal attractiveness, charm or charisma. Why should the amoralist worry about contempt from moralists if he is more than compensated by the admiration of others? There is plenty of empirical evidence that amoralists can be highly successful and lead correspond-

ingly happy lives not in spite of but exactly because of their amoralism and recklessness (cf. Babiak/Hare 2009). They tend to be more ambitious, more risk-seeking and more inspiring than more moral people and thereby able to compete more successfully on the economic, scientific, artistic and political labour market. Since self-esteem is largely dependent on social esteem, even the amoralist's self-esteem need not suffer, provided his own super-ego is sufficiently tolerant. It may well be that he will have to do without many of the options for cooperation offered to less selfish people. But he may be clever enough to act in a way that establishes a reputation of being sufficiently dependable to be given opportunities to achieve his personal aims with the co-operation of others.

Another argument commonly put forward by moralists is hardly more convincing. It says that whoever, as an amoralist, tries to live without commitment to moral norms and values will miss possibilities of experiencing life as meaningful that are, for people with moral commitments, an important source of satisfaction. A life without genuinely moral emotions is, according to this argument, an impoverished life. A considerable part of affective experience – the moral emotions – will be inaccessible to the amoralist. His emotional life will be amputated, as it were, by the absence of emotions such as guilt and remorse, indignation and forgiveness, moral pride and moral self-gratulation. But, again, the amoralist can in many cases easily compensate these limitations in his range of emotions by other kinds of emotions, just as the areligious person is able to compensate the dimension of religious emotions by other emotions. His passions will go in a different direction but they need not be less intense. The egoist's passions will be focused on self-centered emotions like ego-ideals, experiences of success and failure, pride, honour and self-satisfaction. To many artists, scientists and economic leaders ego-ideals and perfectionist ambitions are a far more powerful motivation to give their best than moral ideals. They are productive and creative because they feel a strong urge to live up to images of their own grandeur. Nor can it be doubted that they draw a great deal of satisfaction from being successful in this, especially if their achievements helps them to gain recognition, both from their family, friends and colleagues and from the general public. There is little to suggest that morality is so vitally important for happiness that even a moderately hardened egoist will be persuaded that morality is a necessary condition of well-being.

My conclusion is that not everyone who wants to live as an egoist or amoralist can be given a reason to opt for the moral option in cases of conflict with his or her desires, and that the second horn of the dilemma has proved to lead to an impasse no less than the first. There is, however, an important proviso: This result holds primarily on the level of theory. What it means in practice depends crucially on what might be called the "level of moral aspiration" of a given (so-

cial or private) morality. A morality can be more or less stringent, or more or less demanding. The less stringent it is, the more probable it is that there are no, or very few, situations in which the demands of morality deviate from the demands of self-interest. Given that social conformity and the desire to avoid the disapproval and to gain respect and support from significant others is a strong motive in humans, this motive may well be sufficient to make an average person conform to the moral norms of his or her group, thus harmonizing self-interest and morality. As was shown by Kohlberg and others (cf. Kohlberg/Turiel 1971, p. 440), most people follow the moral norms of their society primarily not for intrinsic reasons but for reasons of social conformity, i.e. in order to avoid what Bentham called "popular sanctions" (Bentham 1948, p. 25). The question, "Why be moral?" does not in general arise for them because the moral thing coincides with what satisfies at least one basic and strong interest. The gap between morality and self-interest will be the broader, however, the more stringent the relevant morality (of an individual, a group or society at large) is (or is believed to be) and the more and heavier sacrifices it demands in the satisfaction of the individual's desires.

It is interesting to see that this scepticism about the chances of giving morality a foundation in rational egoism was shared by one of the great moralists of the nineteenth century, John Stuart Mill – *pace* his attempts to give morality, in *Utilitarianism*, a basis in individualistic hedonism. As his notes on Plato's *Gorgias* show, Mill was well aware of the difficulties of giving morality a basis in individual egoism. In the *Notes*, he presents himself as thoroughly sceptical about the chances to convince the stubborn egoist of the usefulness he can expect for himself from moral virtue. For Mill, ironically, the "sacrifices" appealed to by the moralist as being part and parcel of amoralism are mostly on the side of moralism:

> It is impossible, by any arguments, to prove that a life of obedience to duty is preferable, so far as respects the agent himself, to a life of circumspect and cautious selfishness. It will be answered, perhaps, that virtue is the road to happiness, and that "honesty is the best policy". Of this celebrated maxim, may we not venture to say, once for all, without hesitation or reserve, that it is not true? The whole experience of mankind runs counter to it. The life of a good man or woman is full of unpraised and unrequited sacrifices. (Mill 1978, p. 149)

According to Mill, historical evidence fails to support the hope of moralists that there is a fair balance between morality and happiness, moral desert and destiny. At the same time, Mill is anxious to differentiate. The human tendency to conformism may in fact be sufficient to make people follow what morality requires given that what morality requires is hardly more than the observance of the most fundamental rules necessary to keep society going. As soon as morality aims at

something higher, at the improvement of society or mankind, egoism or amoralism will no longer be supportive of morality. Intrinsic moral motivations are indispensable, at least on the part of an elite with the ability to influence public opinion:

> They, indeed, who have no conception of any higher honesty than is practised by the majority of the society in which they live, are right in considering such honesty as accordant with policy. But how is he indemnified who scruples to do that which his neighbours do without scruple? Where is the reward, in any worldly sense, for heroism? (Mill 1978, p. 149)

Intrinsic motivations are required in order for society to elevate itself above the most elementary moral minimum. The principles of morality must have been successfully internalized by upbringing, education and other social forces in order to make moral virtue a sufficiently stable part of one's identity to motivate unselfish actions. The great achievement that Mill attributes to Plato's several arguments in the *Gorgias* to show that virtue pays even for the convinced egoist is, paradoxically, their utter failure. What they show is that without the preexistence of moral motivations purely rational arguments are unpersuasive. Only those who are already socialized in a morality will happily accept these arguments – not because they are rationally convinced by them but because they lead to conclusions they experience as supporting their own preexistent moral commitments. Only they will be sympathetic to Plato's suggestion that whoever acts contrary to morality acts against the health of his own soul:

> Those only will go along with Socrates in the preceding dialogue, who already feel that the accordance of their lives and inclinations with some scheme of duty is necessary to their comfort; whose feelings of virtue are already so strong, that if they allow any other consideration to prevail over those feelings, they are really conscious that the health of their souls is gone, and that they are, as Plato affirms, in a state of disease. (Mill 1978, 150)

Thus, the second horn is also seen to lead to a kind of circularity. Even the nonmoral justification of the priority of the moral, at least that suggested by Plato, will only be accepted by whoever is already convinced that immorality harms his or her mental integrity. The amoralist has no reason to worry about this kind of ill-health unless he believes in a Last Judgment, karma or any other form of metaphysical sanction.

5 Conclusion

The conclusion to be derived from the above arguments is that the thesis that the Why-be-moral problem is a pseudo-problem is in fact true – on the condition that ultimate irresolvability is a sufficient condition for a problem to be classified as a pseudo-problem. If more is required, namely that this *pseudo* character is so obvious that it makes any effort at closer inspection superfluous from the start, the answer is negative. The irresolvability of the problem is by no means obvious. To that extent at least, the examination of the pros and cons of the pseudo-problem thesis presented in this contribution may not have been futile.

Bibliography

Babiak, Paul/Robert D. Hare, (2006): *Snakes in Suits. When Psychopaths Go to Work.* New York: Harper.

Bayertz, Kurt (2004): *Warum überhaupt moralisch sein?* Munich: C. H. Beck.

Bentham, Jeremy (1948): *An Introduction to the Principles of Morals and Legislation.* New York: Hafner.

Birnbacher, Dieter (1990): "Wittgenstein und die 'Grundfrage der Metaphysik'". In: Rudolf Haller/Johannes Brandl (Eds.): *Wittgenstein - eine Neubewertung. Akten des 14. Internationalen Wittgenstein-Symposiums 1989*, Wien: Hölder-Pichler-Tempsky, pp. 228-238.

Birnbacher, Dieter (2013): "Moral and Other Values". In: *Kultura i Wartosci* 4, pp. 43 – 59.

Blackburn, Simon (1986): *Spreading the Word. Groundings in the Philosophy of Language.* Oxford: Clarendon Press.

Hare, Richard M. (1963): *Freedom and Reason.* Oxford: Clarendon Press.

Kant, Immanuel (1963) *Lectures on Ethics*, trans. Louis Infield. New York: Hackett.

Kant, Immanuel (2002): *Groundwork for the Metaphysics of Morals*, trans. Allen W. Wood. New Haven/London: Yale University Press.

Kohlberg, Lawrence/Elliot Turiel (1971): "Moral Development and Moral Education", In: Gerald S. Lesser (Ed.): *Psychology and Educational Practice.* Glenview IL: Scott/Foresman, pp. 410 – 465.

Mackie, John L. (1977): *Ethics. Inventing Right and Wrong.* Harmondsworth: Penguin.

Mill, John Stuart (1978): "Notes on Some of the More Popular Dialogues of Plato". In: *Collected Works*, vol. XI, Toronto: University of Toronto Press, pp. 37 – 238.

Monro, D. H. (1967): *Empiricism and Ethics.* Cambridge: Cambridge University Press.

Montada, Leo (1993): "Moralische Gefühle". In: Wolfgang Edelstein, Gertrud Nummer-Winkler and Gil Noam (Eds.): *Moral und Person.* Frankfurt am Main: Suhrkamp, pp. 259 – 277.

Nielsen, Kai (1970): "Why Should I Be Moral?" In: Kenneth Pahel/Marvin Schiller (Eds.): *Readings in Contemporary Ethical Theory.* Englewood-Cliffs, NJ: Prentice-Hall, pp. 454 – 483.

Rescher, Nicholas (1984): *The Riddle of Existence.* Lanham, MD: American University Press.

Sen, Amartya (1974): "Choice, Orderings and Morality". In: Stephan Körner (Ed.): *Practical Reason*. Oxford: Blackwell, pp. 54 – 82.

Thornton, J. C. (1970): "Can the Moral Point of View Be Justified?" In: Pahel, Kenneth/Schiller, Marvin (Eds.): *Readings in Contemporary Ethical Theory*. Englewood-Cliffs, NJ: Prentice-Hall, pp. 442 – 453.

Wallace, G./Walker, A. D. M. (Eds.) (1970): *The Definition of Morality*. London: Methuen.

Peter Schaber

Why Be Moral: A Meaningful Question?

In Plato's *Republic* Glaucon tells the story of Gyges, who discovered a golden ring that could make him invisible: "... and becoming aware of this, he immediately managed things so that he became one of the messengers who went up to the king, and on coming there he seduced the king's wife and with her aid set upon the king and slew him and possessed his kingdom. If now there should be two such rings," Glaucon added, "and the just man should put on one and the unjust the other, no one could be found, it would seem, of such adamantine temper as to persevere in justice..."[1] In his comment, Glaucon makes explicit what the story is about, namely, the question that has occupied moral philosophy since its beginnings: Why should we be moral? Why should Gyges, Glaucon asks, if he could make himself invisible and so fear no sanctions for his actions, not seduce the queen, slay the king and take the kingdom? Would it not be irrational to pass up such an opportunity?

Many see the question raised in Glaucon's story as *the* skeptical challenge to moral philosophy: Why should I do what is morally right? Clearly whoever asks this question has a more than superficial interest in ethics. This person wants to know whether he also has reasons to do what is morally right. We should not rule out the possibility that although being moral is undoubtedly a good thing it is by no means a rational thing to do. It might be that in many cases we simply do not have any compelling reason to do what is morally right.

Such a conclusion would be sobering for moral philosophy since it would contradict the widely shared conviction that moral principles are universally binding. This means that a moral principle is considered binding for every person in every situation, unless it conflicts with a stronger moral principle. Whoever wants to hold onto this idea of the universality of moral principles must therefore find a satisfactory answer to the question "Why be moral?" And in fact many moral philosophers do believe it is a central goal of moral philosophy to find a convincing answer to this question.

However, not all philosophers agree. H.A. Prichard, for example, famously argued that the question "Why be moral?" is meaningless because there can be no proper answer. And the reason is that Prichard assumes that there are in principle only two possible answers to the question why we should be moral. Either we should be moral for non-moral reasons or for moral reasons:

1 Plato, Republic, 359 – 360d.

> So far as I can see, the answers all fall, and fall from the necessities of the case, into one of two species. Either they state that we ought to do so and so, because as we see when we fully apprehend the facts, doing so will be for our good, i.e. really, as I would rather say, for our advantage, or, better still, for our happiness; or they state that we ought to do so and so, because something realized either in or by the action is good. In other words, the reason 'why' is stated in terms either of the agent's happiness or of the goodness of something involved in the action.[2]

For Prichard both answers are unsatisfactory. If we give non-moral reasons, then we give the wrong reasons. If we give moral reasons, then the answer presupposes what it is supposed to show, namely that we ought to do what is morally required. As Prichard puts it:

> If the motive in respect of which we think an action good is the sense of obligation, then so far from the sense that we ought to do it being derived from our apprehension of its goodness, our apprehension of its goodness will presuppose the sense that we ought to do it.[3]

It is, if we follow Prichard, impossible to give a proper answer to the moral skeptic.

Is the question "Why be moral?" meaningless then and any further discussion of it futile? As I will argue, Prichard does have a point. The question—"Why should I be moral?"—is, as I will argue, meaningless, if it is understood as the question of whether we have reasons to do what is morally required. That way it is understood as the question why we should comply with moral principles: Do I have reasons to do what is morally required? It is, on the other hand, a meaningful question if it is understood as the question of what the reasons are that justify the claim that we ought to act in certain ways.

This latter question asks for a justification of a moral principle. I will argue that the reasons that justify moral principles are at the same time the reasons why we should obey these moral principles. In other words, the reasons for obeying moral principles are equivalent to the reasons that give these principles their moral justification. If one accepts principles as moral principles one accepts that there are reasons to comply with the principles. As a question about independent reasons for obeying moral principles the question "Why be moral?" does then not make sense anymore. It is devoid of content.

This is not just about the sense or nonsense of asking a question that has been repeatedly and intensively discussed throughout the history of moral philosophy. It is above all about determining what it means to justify moral princi-

2 Prichard (2002), 8.

3 Prichard (2002), 11.

ples. This can mean a number of things, including the attempt to give an answer to the question "Why be moral?" And in fact many moral philosophers would see their work in just this light. But I think this is not the right way to understand justification. Justifying moral principles does not consist in answering the question "Why be moral?" but in answering the question what are the right moral principles. Moral justification is, I believe, about the question of what speaks in favor of a claim that we ought, morally, to act in a certain way. And if this question is answered, there is no further question of whether we have reasons to obey the moral principle. We have reasons to obey moral principles: the reasons which speak in favor of the claim that we ought to act as the principle tells us.

1 The Moral Skeptic

Moral philosophers have provided a variety of answers to the question "Why be moral?" One of them is—as Socrates argued—that acting morally is necessary for living a truly happy life. Another answer is—as for instance Christine Korsgaard thinks—acting morally is constitutive of our identity as human beings.[4] And finally Kant believed that we follow practical reason only by acting morally.[5] I will not discuss these proposals here. I will rather focus on the assumptions that those giving these kinds of answers to the question "Why be moral?" make.

These answers are all addressed to the same audience. They are addressed to the person who asks why he should have anything at all to do with morality, to someone who is "outside" morality and not to someone "inside". This person outside morality can be called a "moral skeptic".[6] And it is this moral skeptic that must be convinced of the binding character of morality, if one wants to show that moral principles are binding for everyone, also for those who have no moral inclinations. If we can convince the moral skeptic that he has to comply with moral demands, then we can convince anyone; that is, if we can show the moral skeptic that he has reasons for following moral principles, then we can show anyone. And this would mean we have proved that moral principles are universally binding. If we have been able to do this, then we have, as many moral philosophers think, justified morality.

4 Cf. Korsgaard (1996), 17.
5 Cf. Kant (1785), 429.
6 Cf., for instance, Stemmer (2000), 17.

If we are to be successful in this task, we must first know more about what kind of person the moral skeptic is, or if there were no such person, what kind of person he would be. This will allow us to determine what question he is asking exactly, and thus what kind of reasons we are looking for. I will follow here Peter Stemmer's account of the moral skeptic. According to Stemmer the moral skeptic has the following three characteristics:

1) The moral skeptic does not have any altruistic predispositions or interests. In other words, he has no interest in showing consideration for others. "It is not part of his conception of himself to be someone who does something for another person."[7] The reasons we have for obeying moral principles cannot then involve empathy or compassion.
2) The moral skeptic is outside morality. According to Stemmer, he does not take any moral standpoint: "What he is asking is why he should take this standpoint."[8] And so he will not be persuaded if we give him moral reasons.
3) Moreover, the moral skeptic rejects the idea that there are objective values and norms. For Stemmer, it is thus "unthinkable in a discussion of the question why he should act morally to refer to the objective existence of moral norms or claims as part of the world."[9]

This is the person we have to convince. But of what exactly should we convince him? He should be convinced of complying with moral principles. What kind of principles are we talking about here? It cannot be about complying with just any kind of principle. We aim at convincing the skeptic that he should comply with *moral* principles. If principles are not justified, no one—be it a skeptic or not—has a reason to comply with them. Therefore, the skeptic cannot be talking about such pseudo-principles but rather is talking about principles that he already acknowledges are morally justified, that is, the right moral principles. The moral skeptic must recognize certain moral principles as the right moral principles. Otherwise the question of whether he should comply with the principles would not be sensible. His question then is: Why should I follow the right moral principles? What are the reasons he has to comply with these principles? This seems to be the question the moral skeptic can sensibly ask.

Here is an interesting distinction to be made. There are reasons why one should comply with a principle. And there are reasons why a principle should be seen as a moral principle. Reasons of the second kind are ones that justify

7 Stemmer (2000), 17 (translated by Paul Lauer).
8 Stemmer (2000), 18.
9 Stemmer (2000), 18.

moral principles. They can be found in theories of moral obligation, for example in utilitarianism, in contractualism, in Kantian ethics and others.

Take a utilitarian who thinks that we ought to give ten per cent of our income to charity. Why—one can ask—should we do so? The utilitarian might say, " because it prevents bad things from happening". This is a justificatory reason for the principle that we ought to give ten per cent of our income to charity. Moral theories such as utilitarianism provide us with justificatory reasons for moral principles. Whether these reasons are really justificatory reasons for moral principles is what the moral debate is all about. Thus, one might reject the given utilitarian justification. One might argue that the fact that bad things can be prevented from happening is not a reason why we ought to act in a certain way. So it is an open question whether the given reason is really doing the justificatory work it is supposed to do.

What is important in our context is this: With regard to any moral principle two different questions can be asked: (a) I can ask whether the principle that we ought to give ten per cent of our income to charity really is a moral principle, that is, something that we really ought to do. Some utilitarians would say yes, some deontologists would deny this. (b) I can ask whether I have any reasons to do what I think I ought to do, for instance, to give ten per cent of my income to charity. The moral skeptic asks—as many moral philosophers think—the second, not the first question. The moral skeptic thinks that certain moral principles are correct. But he would like to know what reasons he has to obey them.

Accordingly, moral theories that intend to answer the skeptic's question proceed in two steps. First, they determine what the right moral principles are; second, they answer the skeptic's question, "Why should I follow these principles?" Thus, one could argue first why, for instance, utilitarian moral principles should be accepted as the right moral principles and then address the question why these utilitarian principles should be followed. This is what John Stuart Mill does in his book *Utilitarianism*. He first tries to provide us with a proof of the utility principle and is then asking why the principles should be followed.[10] The same structure can be found in Peter Singer's *Practical Ethics*. He first argues in favor of certain moral principles, trying to show what we ought, morally, to do with regard to various practical issues: how animals should be treated, what should be done about world poverty etc. In the final chapter 12 then he writes:

10 Mill (1979), 32: "This firm foundation is that of the social feelings of mankind—the desire to be in unity with our fellow creatures, which is already a powerful principle in human nature, and happily one of those which tend to become stronger ... from the influences of advancing civilization ..."

> Previous chapters of this book have discussed what we ought, morally, to do about several practical issues and what we are justified in adopting to achieve our ethical goals. The nature of our conclusions about these issues—the demands they make upon us—raises a further, more fundamental question: why should we act morally?[11]

As concerns the question "Why should we act morally?", things such as a certain idea of a good life, fear of sanctions, or one's own conscience might provide us with the reasons to comply with what has been shown as what we ought, morally, to do. This at least seems to be the idea of how a moral theory should proceed.

2 Stemmer's Model of Justification

Not everyone agrees with this model of moral justification. Stemmer, for instance, rejects it. It is interesting to note why he rejects it.

Stemmer thinks that the question of the moral skeptic is not subordinate to the actual justification of moral principles. On the contrary, he believes that the justification of moral principles is nothing less than answering the question of the moral skeptic, i.e. answering the second question. And this is his argument: Stemmer thinks that the practical reasons a person has depend on the goals that person has. I have a reason to do x, provided doing so furthers my goals.[12] This is also true for those things we ought to do in a moral sense. A person ought to do x, if and only if her goals would call for her to act in this way. Moral obligations are, Stemmer thinks, relative to the goals people pursue and so they are always conditional obligations.[13] A moral principle is then only valid for a person if complying with this principle is in accordance with his own goals. The answer to the question of the moral skeptic does not assume the existence of a list of justified moral principles. On the contrary, such a list can only be drawn up in the light of the goals an individual has.

Is the question of the moral skeptic then also a question about the right moral principles? I think that this is indeed the case, but not for the reasons that moral philosophers such as Stemmer address. I think that the justificatory reasons for moral principles are at the same time the reasons we have for following these principles. Thus, if moral principles were justified by the reasons that are provided by our own goals and by nothing else, Stemmer would be right in

11 Singer (1993), 314.

12 Stemmer's view is reminiscent of David Hume's desire-based view of practical reason according to which reason can only show me how to efficiently fulfill my desires.

13 See Stemmer (2000), 63.

saying that the justification of moral principles would at the same time provide reasons for the skeptic to do what he morally ought to do, but not for the reasons Stemmer claims. It can, of course, be questioned whether the reasons that justify moral principles are the reasons that are provided by the goals people pursue. Such an explanation seems to be implausible because if it were true, we would no longer be able to distinguish moral principles from advice. If these principles were only valid when they served an individual's goals, then the moral principle would take the form of "Since you have this or that goal, then my advice to you is to do this or that." And we would have to interpret statements such as "You should do this or that" as advice. But then it would no longer be clear whether we could still meaningfully use moral language and, for example, be justified in calling some actions morally wrong. Whoever does not follow advice may be foolish or even stupid, but he is not acting immorally.

3 Arguing within Moral Theory

What is right about Stemmer's argumentation is that the justification of morality does not take two steps but only one. The justification of moral principles provides us with the reasons we have to obey these principles. Say, for instance, that we ought to help people in need. The reason why this principle holds is—whatever it turns out to be—the reason why we should act according to it. If one accepts that it is a moral principle that we ought to help people in need, then one accepts that there are reasons why this is so, i. e. that there are reasons for compliance. If this were not the case, it would not be seen as a moral principle in the first place. And the moral skeptic would not have any reason to ask why he should comply with the principle, for since he grants that it is a moral principle he also grants that there are reasons to comply with it. Thus, if one asks the question "Why be moral?" one cannot sensibly ask whether there are reasons to comply with the moral principle that we ought to help people in need, for instance.

One could say that the moral skeptic is asking whether the reasons that justify a moral principle are at the same time reasons for him. He could assume that there are reasons why one ought to help people in need, at the same time however believe that he need not himself follow that principle. Then the question he is asking would be: Are the reasons that justify the moral principle reasons for him? But this question does not make any sense. If the skeptic accepts a principle as a moral principle ("we ought to help people in need"), he at the same time accepts that there are reasons for him that justify the principle in question (without necessarily knowing them). If this were not the case he would not accept the

principle as a moral principle. He might very well do so, but then his doubts could not be understood as doubts about the idea that he has reasons to do what he accepts is morally required. If the skeptic asks for the reasons to do something that is demanded of him, then he can only sensibly ask whether this demand is justified. This is however not the question why he should be moral but the question whether or not this principle is a member of the class of morally right principles. But this would mean that the moral skeptic is no longer arguing from a standpoint outside morality but from one inside. He is taking part in a debate about the content of morality and not about the reasons for obeying moral norms. If the moral skeptic on the other hand acknowledges that the principles he is being obligated to obey are the right moral principles, then it is indeed unclear what he is asking about. Whoever says, "Yes, those are the right moral principles", and then asks whether he should obey them is asking a question that does not make any sense.[14]

4 Reasons for the Moral Skeptic

Whoever shares the widely held view that moral principles are universally binding must show that even a moral skeptic has reasons to accept the moral principles. And if what I have said so far holds, then the reasons the moral skeptic has to accept moral principles are nothing other than the reasons that justify moral principles. Thus, he might have to accept that he ought to help people in need because this would save them from very bad things.

Will the moral skeptic who is outside morality be persuaded by such reasons? Well, it could be that the moral skeptic will not enter the realm of reasons, that is, he will rule out reasons in general terms. If that is so, then he will remain unpersuaded. That does not however mean that there are no reasons for him to act morally. It simply means that he refuses to consider such reasons. The fact that there are reasons to act morally—for example, to help others in need—is unaffected by such a refusal. The moral skeptic could however simply refuse to accept that something given as a reason (for example, that very bad things should be prevented) is a reason. In this case we would have to discuss what reasons justify moral principles. This is what moral philosophers do. However, this is not a debate—and this is important in our context—that takes place outside mor-

14 If one accepts moral principles, then one necessarily accepts that one has reasons to act accordingly. One might, of course, have at the same time reasons not to act accordingly. Reasons of self-interest might speak against acting morally. Still, one has reasons to obey the moral principles.

ality. It is nothing other than what moral philosophers discuss. They are arguing about justifications of moral principles: what we morally ought to do and why we ought to do such things. Can the moral skeptic be convinced why he should do so?

Let us assume that the principle "we ought to help people in need" is correct. And the reason is that by doing so very bad things can be prevented. This is what we tell the moral skeptic. The moral skeptic might not be convinced. He doubts that these are justificatory reasons for the principle in question. Thus, he doubts whether the principle that "we ought to help people in need" is a moral principle. If he does so, he takes part in the ordinary moral debate on the right principles of morality. Or he might accept the principle: "Yes, we ought to help people in need". And he might then ask why he should to do so. But what then is he asking for? He is not asking whether there are any justificatory reasons for the principle, because this he already accepted by accepting the principle as a moral principle. One might say that he is asking the question of whether following the principles is serving his own interests. This is, of course, a sensible question. The moral skeptic can ask whether acting morally serves his interests. But does the question "Why be moral?" address this concern? The justificatory reasons of moral principles are the reasons why we should be moral. And they are accepted, if the principles themselves are accepted as moral principles. So the moral skeptic is not asking for those reasons, provided he has accepted a principle as a moral principle. And if someone asked: "Are there reasons to follow these principles?" the skeptic would say "Yes, there are: the reasons which justify the principle". These are at the same time the reasons to obey the principle. Of course, he could ask whether following the principles is serving his own interests. They might differ, they might not. If the moral reasons differ from the reasons of self-interest, the moral skeptic would still have reason to follow the moral principles. The justificatory reasons would still obtain. He might at the same time have no reasons of self-interest to follow the moral principles. Some think that this implies that he had no reasons to act morally. But this is wrong. He has reasons to act morally, namely those that justify the moral principles.

5 The Answer to Glaucon's Question

I began this paper by reminding us of Glaucon's story about the shepherd Gyges, and I would like to return to it. Why should Gyges—this is Glaucon's question—not have seduced the queen, slain the king and taken the kingdom since he would have been able to make himself invisible and would not have had to

fear any reprisal? Would it not have been stupid for him to have passed up this opportunity? How should we respond to Glaucon? I think that we should give him our moral reasons for not acting like this, for example, tell him that we should not act like this because it violates the fundamental right of another person to life.

Of course Glaucon could ask why it is wrong to violate a fundamental right of another person. He would, however, simply doubt that violating a fundamental right of another person is a reason not to act in a certain way. We find ourselves then once again discussing the content of moral reasons. Can the moral principle not to treat the king the way that Gyges did in the story be justified with reference to a fundamental right or not? Whatever the answer to this question is, it is also—as I have tried to show—an answer to Glaucon's question "Why be moral?" This does not mean that Gyges necessarily would have acted differently if he had acknowledged that the fundamental rights of the king are violated by deception and murder. When we talk about reasons for action, then we are not *eo ipso* talking about what motivates people but merely about what should motivate people. The question "Why be moral?" is, however, not a question about the actual motivation of people but a question about reasons for accepting principles as moral principles. As far as these reasons are concerned, moral philosophy does not need to make any special effort. The reasons we have in mind when we say that we must act in certain ways are reasons that we should all act on. And the reasons not to do what Gyges was said to have done are nothing but the reasons that justifiably demand that he refrain from such actions. There is no further question about reasons. Whoever knows he ought, morally, to do something knows that he has reasons to act accordingly. Gyges might not recognize that he ought, morally, to do something. He might reject that there are any reasons to act morally. And if he then asked why he should be moral, he might not be sure whether there are really no moral principles ("Are there reasons to accept moral principles I have not yet discovered?"). But once he accepts moral principles the question "Why be moral?" would be meaningless.

It may seem as if I have argued in favor of H.A. Prichard's claim that the question "Why be moral?" has no meaning. But I do not think that this is the case. The question is meaningless if it is understood as the question of whether there are any reasons to comply with moral principles. But the question is a sensible one, if it is understood as the question of whether there are reasons to accept moral principles in the first place. If I think that we ought to help people in need, I cannot sensibly ask whether I have reasons to do so. But I could have doubts whether we morally ought to do help people in need. And I might ask whether we ought to help people in need. This is a sensible first order moral question. It needs good moral theorizing to answer it.

Bibliography

Kant, I. (1785): *Grundlegung zur Metaphysik der Sitten*, translated by L. White Beck (1990) as *The Foundations of the Metaphysics of Morals,* Second Edition, New York: Macmillan.
Korsgaard, Ch. (1996): *The Sources of Normativity*, Cambridge: Cambridge University Press.
Mill, J.S. (1979): *Utilitarianism*, ed. by G. Sher, Indianapolis: Hackett.
Plato (1961): *The Collected Dialogues of Plato including the Letters*, ed. by E. Hamilton and H. Cairns, Princeton: Princeton University Press.
Prichard, H. A. (2002): "Does Moral Philosophy Rest On a Mistake?", in *Moral Writings*, Oxford: Oxford University Press, pp. 7–20.
Singer, P. (1993): *Practical Ethics*, Cambridge: Cambridge University Press.
Stemmer, P. (2000): *Handeln zugunsten anderer*, Berlin/New York: Walter de Gruyter.

Part II: **Exploring Viable Answers**

Robert B. Louden

Why Be Moral? A New Answer to an Old Question

The Ultimate Question?

The question, "Why Be Moral?" is a staple in many anthologies of ethical theory[1] as well as in numerous textbooks[2] and monographs[3] on ethics. It has also been described variously as "the supreme question" (Stace 1937, p. 252), "the Ultimate Question" (Taylor 1978, p. 483),[4] "the profoundest problem of Ethics" (Sidgwick 1907, p. 386n.4),[5] and "the Holy Grail of moral philosophy" (Hills 2010, p. 3).[6] Is there anything new to say on this topic? I believe there is. In this essay I wish to offer a new and more plausible answer to this venerable question, primarily because all of the existing answers with which I am familiar appear to be unsatisfactory.

I will say more shortly about why the existing answers seem to be unsatisfactory, but first I wish to say a few words about the way in which the question is understood in most philosophical discussions. There is a clear sense in which the correct answer to the question, "Why Be Moral?" is simply "because it's right" (Hospers 1961, pp. 191–98), and this is one reason why F. H. Bradley begins his classic essay, "Why Should I Be Moral?" by observing: "Why should I be moral? The question is natural, and yet seems strange. It appears to be one we ought to ask, and yet we feel, when we ask it, that we are wholly removed from the moral point of view" (Bradley 1927, p. 58). I believe that Bradley is right: when we try to find an answer to the question, "Why be moral?" we are, as he puts it, "wholly removed from the moral point of view." This is also one rea-

1 Cf. Hospers and Sellars 1970, Pahel and Schiller 1970, Pojman 1995, Shafer-Landau 2007, Sher 2012.

2 Cf. Hospers 1961, Pojman 1990, Shafer-Landau 2010.

3 Cf. Bradley 1927, Stace 1937, Toulmin 1950, Frankena 1980, Gert 1988.

4 However, if one is searching for a really ultimate question, the case that Camus makes for his own candidate does seem rather compelling: "There is but one truly serious philosophical problem, and that is suicide. Judging whether life is or is not worth living amounts to answering the fundamental question of philosophy. All the rest – whether or not the world has three dimensions, whether the mind has nine or twelve categories – comes afterwards. These are games; one must first answer" (Camus 1955, p. 3).

5 See also Parfit 2011, 1: pp. 141–49.

6 See also Blackburn 1984, p. 222.

son why Prichard attacks the very attempt to give arguments for being moral, holding instead "that we do not come to appreciate an obligation by an *argument*, i.e. by a process of non-moral thinking" (Prichard 1968, p. 9). People who are already committed to the moral point of view do not need any additional reasons to be moral, and the question, "Why be moral?" will appear strange to them. If they view a certain action as being morally right, then they believe they ought to do it – its rightness already supplies them with a sufficient reason for performing the act (see Hospers 1961, p. 193). But the "Why Be Moral?" question, as interpreted by most philosophers, is understood as a question asked by someone who is *not* already committed to morality. It is a skeptical question posed by someone who stands outside of morality, not one asked by someone who already subscribes to the moral point of view. The question is normally interpreted as being asked from outside a moral framework, and thus translates into something like "Why should I step inside a moral framework?" – where the 'why' is shorthand for "are there any compelling reasons?" Are there any compelling reasons for me to be moral, or should I simply remain outside of morality? Why is it rational for me to be moral? In effect, we are looking for nonmoral reasons to be moral, and then asking if any of them are strong enough to lead us to give up our skepticism about morality.[7]

Two related points: 1) Some philosophers have held that it is impossible to find nonmoral reasons that would convince a moral skeptic to adopt the moral point of view. For instance, Prichard's famous claim that moral philosophy "rests on a mistake" (Prichard 1968, p. 1; cf. Gert 1988, pp. 228–29) – viz., the belief that a rational proof can be given to show people that they ought to be moral, has often been interpreted in this manner. On this view, the question "Why be moral?" does not admit to an answer and is in effect closed. Morality is viewed as a kind of exclusive club – if you're not already a member, there is no way to be admitted. But I view it as an open question: we don't know if a convincing answer can be given to the moral skeptic until we try. 2) It was popular for a while among analytic philosophers to argue that the question "Why be moral?" is simply not a meaningful question but only a pseudo-question (Thorn-

7 The unsatisfactoriness of Bernard Gert's answer to the question "Why be moral?" should now be apparent. He writes: "one should be moral because he will cause or increase the likelihood of someone suffering evil if he is not" (Gert 1988, p. 228; see also p. 234). But as he himself notes, "this is a moral reason or answer to the question 'Why should one be moral?'" (p. 228). However, what we need is a nonmoral reason. Gert's answer may work for someone who is already committed to morality, but it won't work for someone who stands outside of morality. In effect, he's saying "one should be moral because it's morally wrong not to be moral." His answer is question-begging.

ton 1970, p. 453), like so many traditional philosophical queries whose absurdities needed to be exposed in order to advance knowledge. The absurdity of the question "Why should I be moral?" became apparent, it was claimed, once the implied 'should' ("Why *should* I be moral?") is construed as a moral 'should'. Interpreted in this manner, the question then becomes vacuous, like asking "Why are all round things circular?" (Nielsen 1970, p. 748; cf. 1958) or "are there some emerald objects that aren't green?" (Toulmin 1950, p. 162). However, if the 'should' is not interpreted as a moral 'should', then the question is by no means absurd. And as noted earlier, this is in fact the way that the question has traditionally been understood. So we should construe "Why be moral?" as a meaningful question. But – judging by the growing volumes of literature devoted to the problem – it is certainly not an easy question to answer in a convincing manner.[8]

Problems with Previous Answers

I remarked earlier that the traditional answers to the question "Why be moral?" are unsatisfactory. Although I can't examine all of the answers in this short essay, let me try now to briefly sketch what is wrong with at least the most popular ones. From the beginning and up to the present, the most frequently given answer to the question "Why should I be moral?" is "because it is in my self-interest." In order to show that it is rational to be moral, it has been assumed that we must show that it is in one's self-interest to be moral (see, e.g., Gauthier 1970). Let us quickly review some of the relevant points in the frequent appeals to self-interest.

Plato, in Book II of the *Republic* – still today the *locus classicus* of the discussion[9] – has Socrates respond to Glaucon's challenge by arguing that justice

8 Like other traditional philosophical problems that were once dismissed as pseudo-questions (including the question of meaning in life – see Quinn 2008, p. 35), "Why be moral?" has recently gained a new lease on life. But, again, it is not an easy life, for there is little consensus about how to properly answer the question.

9 Cf. Shafer-Landau 2007, pp. 146–52, Sher 201, pp. 7–13, Pojman 1995, pp. 534–40. William Frankena, in his lecture, "Why Be Moral?" pushes the issue back even further when he remarks that "the question is at least as old as Hesiod" (Frankena 1980, p. 78). In *Works and Days* 270–73, for instance, the poet says that he is just because it pays:

> ... I would not be just myself
> nor would I have my son so – for it is a bad thing
> to be just if the unjust should get more justice than the just man;
> but I do not believe that yet Zeus of Counsel will make such an ending (Nelson 1998, p. 16).

in fact benefits its possessors (Plato 1996, p. 1007/367d). In other words, it is to one's own advantage to be moral. Socrates develops his case later in the *Republic* chiefly by arguing that injustice or immorality corrupts one's soul. Immorality corrupts one's soul, it is not in one's interest to have a corrupt soul, so therefore one has a self-interested reason to be moral.

Hobbes, another frequently-cited source,[10] also holds that morality and self-interest are not in conflict, but his focus is on human beings' collective self-interest. It is in our collective self-interest to be at least minimally moral because if we are not, society will fall apart. Hobbesian morality functions as a mechanism of social control, and outside of society there is no morality. Unless and until there is a general adherence to the requirements of morality, human life will once again (as it was before individuals left the state of nature) become "solitary, poor, nasty, brutish, and short" (Hobbes 1946, p. 82). However, as others have pointed out (Hospers 1970, p. 743; Nielsen 1970, p. 751), this Hobbesian strategy, updated by Baier (1958), Kavka (1984) and others in the late twentieth century, though it may provide a plausible answer to the question "Why should people *in general* be moral?" does not address the harder question, "Why should *I* be moral?" For my interests may not coincide with other people's interests. There may be good reasons for people in general to be moral that don't apply to me, particularly if I can successfully obtain what I want without succumbing to the moral constraints adhered to by others.

The traditional religious answer to our question – "I should be moral because God will reward me if I am, and punish me if I am not" – is not discussed much by contemporary philosophers, even though this "divine command" response continues to be the most popular answer to the question among laypeople. But the main point I wish to stress here is that this kind of religious answer to the question, "Why be moral?" is also an appeal to self-interest. As Hospers remarks,

> It is just self-interest pushed into the next world instead of being confined to this one. The person who acts from this motive is just playing the game for higher stakes. He is declaring his willingness to postpone his reward a bit longer in order to collect at a higher rate of interest in the next world – which is as selfish as it could be. It is like working longer hours in order to collect time and a half at the end of the day. (Hospers 1970, p. 738)

These well-known self-interested answers to the question "Why be moral?" can easily be multiplied (and in textbooks they usually are, by adding additional readings on psychological egoism and so-called "ethical egoism"), but they all

10 Cf. Shafer-Landau 2007, pp. 606–19, Sher 2012, pp. 23–33, Pojman 1995, pp. 53–61.

seem to me to be beside the point for the simple reason that morality is seldom about self-interest. Morality often asks us to overcome our self-interest and take a genuine concern for others, and the pull of self-interest itself is the biggest barrier to acting morally. Indeed, even in cases where people appear on the surface to be acting morally, they are usually only acting out of self-interest. This is Kant's point in the *Groundwork* when he remarks

> From love of humanity I am willing to concede that the majority of our actions are in accord with duty, but if we look more closely at the intentions and aspirations in them, we everywhere come up against the dear self [*das liebe Selbst*],[11] which is always flashing forth, and it is on this – and not on the strict command of duty, which in many cases would require self-denial [*Selbstverleugnung*] – that their purpose is based (Kant 1996, p. 62/4: 407).[12]

If the requirements of morality are often inconsistent with the dictates of self-interest, then the attempt to convince someone to be moral by appealing to her self-interest is wrongheaded and essentially irrelevant. For in appealing to self-interest, we are not necessarily making contact with morality at all.[13] Morality and self-interest are not the same – they do not always coincide. Humans are capable of at least occasionally overcoming egoism, and this is what morality frequently asks of them. Here too, my own view on this topic is thus close to Prichard's, who held that there is no reason to suppose that the dictates of morality and self-interest coincide. Much of the history of moral philosophy, he observed, "is occupied with attempts either to prove that there is a necessary connexion between duty and interest or in certain cases even to exhibit the connexion as something self-evident." But his own position, and mine as well, is that these

11 The title of Alison Hills's book, *The Beloved Self: Morality and the Challenge from Egoism*, stems from this passage. However, she uses Jonathan Bennett's translation of the passage, in which "*das liebe Selbst*" is rendered as "the beloved self" (see Hills 2011, p. 4).

12 However, there is a (very) different sense in which morality on Kant's view is consistent with self-interest. For he holds that duties to oneself are "the supreme condition and *principium* of all morality" (Kant 1997, p. 125/27: 344). If there were no duties to oneself, "then there would be no duties whatsoever (Kant 1996, p. 543/6: 417). The work that one needs to do on oneself in order to become an effective moral agent ("normative self-government," as Korsgaard puts it – 2006, p. 112) thus grounds morality. For discussion, see Louden 2011, pp. 17 – 19, 36 – 37. And Kant's appeal to self-respect as a feeling that "is the basis of certain duties, that is, of certain actions that are consistent with his duty to himself" (Kant 1996, p. 531/6: 403) has also been invoked by some as a Kantian answer to the question, "Why be moral?" For discussion of this latter point, see Ricken 2010, pp. 11 – 20.

13 As H.L.A. Hart remarks, "obligations and duties are thought of as characteristically involving sacrifice or renunciation, and the standing possibility of conflict between obligation or duty and interest is, in all societies, among the truisms of both the lawyer and the moralist" (Hart 1961, p. 85).

attempts are "out of place" and should "never have been made" (Prichard 1968, p. 204).

Closely related to the issue of self-interest is the appeal to happiness: "You should be moral because you will be happier if you are." And this answer too goes back at least as far as Plato. In the *Republic*, Socrates replies to Glaucon's challenge by trying to convince readers that a man is truly happy when he is moral. But this alleged connection between morality and happiness quickly begins to vanish upon examination. Being moral certainly is not sufficient for happiness – this was Aristotle's point when he observed that "those people who claim that somebody being tortured on the wheel or meeting with great misfortune is happy – if he is good – make no sense" (Aristotle 1999, p. 117/VII.13 1153b20 – 22). In making his own case for the necessity of "goods residing in the body as well as external goods and chance" (Aristotle 1999, p. 116/1153b18), Aristotle undoubtedly went overboard (as Stoics and Christians would later point out) in claiming that wealth, political power, good children, and personal beauty were all necessary for happiness (see Aristotle 1999, p. 11/I.8 1099a30 – 1099b9), but freedom from constant and excruciating pain does not seem too much to ask for.

But is morality even necessary for happiness? Even this weaker claim is suspect, and both real life and fiction seem intent on knocking it down. Justice does not always triumph, good deeds sometimes do go unrewarded, crime sometimes does pay, and moral virtue does not always make its possessors happy. Nietzsche surely exaggerated when he remarked in *Twilight of the Idols* that "only the Englishman" strives after happiness (Nietzsche 1988, vol. 6, p. 61/I.12) – many people elsewhere also strive after it. But I take his point to be, at least in part, that the English utilitarians are simply wrong in thinking that morality aims primarily at happiness. In real life as well as in many novels and films, there are some people who manage to achieve happiness without being moral, and there are some morally good people who are not happy. Henry Fielding, in his novel *Tom Jones*, puts it well when he observes: "There are a set of religious, or rather moral writers, who teach that virtue is the certain road to happiness, and vice to misery, in this world. A very wholesome and comfortable doctrine, and to which we have but one objection, namely, that it is not true" (Fielding 1950, p. 690).[14] And Leslie Stephen echoes Fielding's remark in the Conclusion to his popular late nineteenth-century textbook, *The Science of Ethics:*

> one great difficulty must remain unsolved. Rather, I assert that it is intrinsically insoluble. There is no absolute coincidence between virtue and happiness. I cannot prove that it is

14 Hospers also cites this passage in his "Why Be Moral?" (1970, pp. 736 – 37).

> always prudent to act rightly or that it is always happiest to be virtuous. My inability to prove these propositions arises, as I hold, from the fact that they are not true (1907, p. 418).[15]

Although the above is by no means intended as an exhaustive examination of all previous answers to the question, "Why be moral?" the main territory has been surveyed, and the central weaknesses of each position have been indicated. What I believe is needed in order to make headway in this ancient and ongoing debate is to locate some nonmoral reasons that are not simply appeals to self-interest or happiness, but yet are powerful enough to convince moral skeptics that they do indeed have sufficient reason to give up their skepticism about morality and instead try to be moral. In the next section of my paper, I will indicate what these reasons are.

The Ethics of Rational Discussion

My strategy for answering the question, "Why be moral?" is in part an application of previous work of mine to a different domain,[16] and it involves drawing attention to the norms and values that lie behind or are presupposed by ordinary cognitive activity and rational communication. There are norms regarding how we ought to think and reason, and some of these norms are arguably moral norms. If we can show that anyone who wishes to successfully engage in such activities must first adhere to certain underlying norms and values (and if we can also show that these norms and values are at least arguably *moral* norms and values), then we have also shown, in effect, that anyone who wishes to engage in ordinary cognitive activity and rational communication must also try to "be moral" – at least in some minimal sense. After presenting my argument, I will also try to show that this strategy avoids the pitfalls associated with earlier canonical answers to the question, "Why be moral?" That is, my answer to the question does not itself involve an appeal to moral reasons (and thus is not a question-begging response), nor does it rest on appeals to self-interest or happiness. Finally, in the concluding sections of my paper, I will also respond to some objections to my position.

What is necessarily presupposed whenever anyone engages in a rational discussion with others? Among other things, the following: honesty regarding facts

15 Frankena (1980, p. 84) and Pojman (1995, p. 531) also cite this passage in their discussions.

16 See Louden 1992, pp. 66–81, Louden 2012, and Louden 2013.

and evidence, respect for the equal rights of others to participate in the discussion, and absence of coercion (refraining from violence) in attempting to justify one's position to others. These necessary assumptions are what Habermas and other discourse ethicists call "presuppositions of communication (*Kommunikationsvoraussetzungen*)" (Habermas 1991, p. 133), and, as Alan Gewirth notes, they are "moral aspects of the intellectual enterprise" (Gewirth 1978, p. 360). These moral aspects[17] of rational discussion and communication lie behind and are presupposed by our cognitive activities, and adherence to them is necessary in order for the activities to function in their normal intended manner. For instance, if (as occasionally happens in real life) someone distorts the data to fit his own personal agenda, or deceives others by making false statements, or uses coercion to force others to accept his position, the discussion itself suffers in a fundamental way – its rational credentials are severely compromised or (in severe cases) forfeited entirely.

Thus far I have presented my position by referring to some underlying norms and values that are presupposed by rational communication and ordinary cognitive activities. But the basic idea can be pushed back still further, to an even more fundamental notion. The slogan "meaning is normative" has been in the philosophical air for some time (cf. Gibbard 2012, p. 6), and in the work of many contemporary philosophers it clearly "plays a pivotal role" (Boghossian 2005, p. 205). Recently, Allan Gibbard has published an entire book devoted to defending a version of the slogan. Earlier defenders of one or another version of the claim that meaning is normative include Saul Kripke (1982, p. 37) (who in turn arrives at it via Wittgenstein – see Wittgenstein 1958, esp. p. 82/§201)[18] and John McDowell (1984, p. 336), but Gibbard traces the key idea behind his approach to following opening statement from Lance and O'Leary-Hawthorne's 1997 book, *The Grammar of Meaning:*

> the very speech act of making a meaning claim is itself normative, ... saying what something means is *pres*cribing. As such, meaning claims have more in common with the claims of morality than they do with the claims of science and so ought to raise for us philosoph-

17 Obviously, much hinges here on what is exactly meant by "moral," and I don't intend to beg the question. See "Moral/Nonmoral," below, for attempt at an answer.

18 Cf. Glock, who argues that "within analytic philosophy, the idea that there is an irreducibly normative dimension to linguistic meaning goes back to Wittgenstein" (Glock 2005, p. 220). But he traces the claim back to pre-*Investigations* Wittgenstein, citing a text of Friedrich Waismann's "from the early thirties which is based on dictations by Wittgenstein" (Glock 2005, p. 220) in which Waismann records Wittgenstein as saying that one needs to regard language "from a *normative* perspective" (Waismann 1976, p. 191, see also p. 185).

> ical questions consonant with those arising in moral philosophy, rather than in philosophy of science. (Lance and O'Leary-Hawthorne 1997, p. 2; cf. Gibbard 2012, p. ix)

Philosophical arguments about "the meaning of meaning" (cf. Ogden and Richards 1923) can and do get complicated quickly, but the basic idea behind Gibbard's version of the thesis that meaning is normative is actually rather simple. As Christopher Hill writes in his review of Gibbard's book, "Gibbard's main claim is that linguistic meaning is normative. ... [He] proposes that claims about meaning imply claims of obligation – more specifically, claims about what we *ought* to say in response to queries, and claims to the effect that we *ought* to accept and reject sentences in various circumstances" (Hill 2013, p. 1). Or, as Gibbard himself puts it, "the slogan 'Meaning is normative' ... cashes out in another slogan, '*Means* implies *ought*'. Ascriptions of meaning imply straight ought ascriptions" (Gibbard 2012, p. 11; cf. 10).

I take the claim that meaning is normative to be consistent with, and at bottom quite similar to, my earlier claim that ordinary cognitive activity and rational communication involve adherence to norms and values. In both cases, we are pointing to a normative dimension that is necessarily involved with something that many people regard as entirely nonnormative. But there is also a sense in which the "meaning is normative" thesis is more radical than my earlier claim. For it implies that one needn't be first engaged in an argument or discussion with anyone else in order to feel the force of an ought. As soon as one commits oneself to the meaning of a linguistic term ("I believe that the concept x means y"), normative commitments are involved. For instance (this is one of Gibbard's examples), if you believe that snow is white, then you *ought* not to believe that nothing is white (Gibbard 2012, p. 13). Similarly, if you believe that cats are mammals, then you *ought* not to believe that there are no mammals with soft fur, a short snout, and retractable claws. For each meaningful linguistic term that a speaker subscribes to, there will be a huge – probably infinite – number of relevant obligations concerning what the speaker ought to believe that follow from her view of what the term means. And this is at least part of what Kripke has in mind when he claims that "the relation of meaning and intention to future action is *normative*, not *descriptive*" (Kripke 1982, p. 37).

I have tried to show that much more is at stake in one's response to the question, "Why be moral?" than has been traditionally assumed. For once we realize how many mundane and fundamental human activities themselves involve adherence to basic norms and values (assuming we can also convince ourselves that these norms and values are arguably *moral* norms and values – see "Moral/Nonmoral," below), we see that the question actually ties in to a wide range of human activities and not merely to "moral conduct," narrowly con-

strued. And this realization should be powerful enough to persuade even the most stubborn moral skeptics that they too should try to be moral. For if they do not, they will not be able to make meaningful claims or engage in rational discussions. Granted, this realization will not necessarily motivate them to strive to become moral saints. But that is a much different problem. The question we are trying to answer is not "Why be as morally good as possible?" but the more modest question, "Why be moral?"

One might respond to the line of argumentation presented thus far as follows: "But aren't you too merely appealing to self-interest and/or happiness? That is, isn't it in one's self-interest to speak meaningfully (make oneself understood), communicate rationally with others, etc.? And aren't people more likely to be happy if they say something coherent and meaningful, communicate rationally with others, etc.? And didn't you argue earlier that appeals to self-interest and happiness are not appropriate responses to the question, 'Why be moral?'"

My response to this objection is "yes and no." Yes, in the sense that I agree both that it is generally in one's self-interest to speak meaningfully and communicate rationally, and that one will often be happier if one acts in these ways. But no, in the (more important) sense that what I have drawn attention to is something much bigger than self-interest and happiness. For instance, even the most selfless altruist imaginable occasionally needs to speak meaningfully and communicate rationally, and when she does so, self-interest will not be her motive – for she is a selfless altruist. Similarly, an extremely disconsolate person – say, Kant's example of someone "who feels weary of life because of a series of ills that has grown to the point of helplessness" (Kant 1996a, p. 73/4: 421) and who as a result is suicidal – even this person will need to at least occasionally speak meaningfully and communicate rationally. And when he does so, he is not aiming at happiness, for he is weary of life. In short, *everyone* – even those who do not act out of self-interest or a desire to be happy – at least occasionally thinks coherently, speaks meaningfully, and communicates rationally. And when they do so, they are forced to also adhere to the underlying norms and values implied by meaning and rational communication. In other words, there are norms and values that govern and permeate human thought, and those who forswear them completely cannot think coherently. And why would people choose to think incoherently?[19]

19 For the purpose of deceiving themselves, one might reply. Possibly. But if they wish to be successful in achieving this aim, they too will need to at least occasionally adhere to the norms and values to which I have drawn attention. For a rationally executed plan of self-decep-

Finally, this argument is also not an appeal to *moral* reasons (e. g., "you should be moral because you'll cause an increase in suffering if you don't"), and thus it is not question-begging. We have presented moral skeptics with powerful nonmoral reasons to be moral, and in doing so we have shown them why they need to step inside what is arguably a moral framework.

A New Answer?

I have billed my response to the question, "Why be moral?" as a *new* answer, but it is time to confess that there is a sense in which it is not new at all. For the earliest and most famous antecedent of the view that I am defending is of course Plato's account of the form of the good in the *Republic*. When Plato's Socrates asserts that "what gives truth to the things known and the power to know to the knower is the idea of the good [*ten tou agathou idean*]" and that the idea or form of the good is "the cause [*aitian*] of knowledge and truth" (Plato 1996, p. 1129/508e), he too is claiming that adherence to moral values is necessarily presupposed by all cognitively meaningful activity. Values, according to Plato, come before facts and make facts possible. As Julia Annas remarks in her exposition of Plato's form of the good, "values are fundamental to explaining facts," they "can be *better* known than facts," "they are *more* fundamental to our understanding" than facts, and "in some obscure way" they are even "*more* real" than facts (Annas 1981, pp. 246, 247).

Much of the contemporary literature that I have drawn from focuses primarily on norms and normativity, whereas Plato's focus is on values and evaluation. The normative domain is centered on concepts such as *ought*, *obligation*, and *a reason*. As Gibbard remarks (with some help from Wilfrid Sellars), "normative judgments are 'ought' judgments. They are judgments that are ... 'fraught with ought'. They are judgments that move within 'the space of reasons'. They are 'oughty'" (Gibbard 2012, p. 10; cf. Millikan 2005, p. 61, Sellars 1956, §36). The evaluative realm, on the other hand, focuses on concepts such as *good* and *bad*, as well as on thicker virtue and vice concepts such as *courageous*/*cowardly* and *generous*/*stingy*.

The normative commitments implied by meaning and rational activity are perhaps easier to see than the evaluative commitments. For instance, as noted earlier, a claim about what a term means also implies claims about how we

---

tion of the sort envisaged ("in order to deceive myself, I will think incoherently") itself implies that the agents in question are consistently following a hypothetical imperative.

ought to respond to queries involving the term, and engaging in a rational discussion implies that the participants *ought* to look at the evidence impartially, that they *ought* to be honest, *ought* to consider counterarguments, etc. But I follow Plato in asserting that there are also value commitments implied in these activities. Ironically, Nietzsche – not exactly a fan of Plato's ("Plato is boring" and "a coward in the face of reality," a symptom "of decay" whose moralism is "pathologically conditioned")[20] – offers some insight in grasping this claim. Behind logic and science, he asserts in *Beyond Good and Evil* and elsewhere, "stand valuations [*Wertschätzungen*] ... For example, that the definite should be worth more than the indefinite, appearance worth less than 'truth'."[21] Nietzsche of course also criticizes Plato's form of the good, but his claim that logic and science presuppose a commitment to the view that truth is good and error bad is nevertheless at least a hint in the direction of Plato's mysterious doctrine that the good is the ultimate cause of the knowledge, truth, and reality of everything else. An additional complication is that in real life the distinction between norms and values is not always rigid – norms and values frequently entail one another. For instance, teleological ethical theories posit a vision of the good (viz., a value) to strive for. But they then attach an *ought* (viz., a norm) to this value: "you ought to realize the good, your duty is to maximize the good."

But my main point regarding Plato at present is that he does not employ his famous doctrine of the form of the good as a response to Glaucon's challenge, "Why be moral?" Rather, the primary role of this doctrine in the *Republic* is to justify the claim that philosophers should rule in a just state. True philosophers have knowledge of the forms, including the highest and most important form of all – viz., the form of the good. And those who truly possess a knowledge of what is good should rule over the rest of us.

Why does Plato not have Socrates employ the doctrine of the form of the good as a way of answering Glaucon's challenge, and instead try to convince Glaucon that it is in his interest to be moral because he will be happier if he is? I'm not sure. But if he had done so, he would have articulated something closer to my view – viz., that one should be moral because one cannot even make meaningful statements or engage in rational communication unless one is at least minimally moral. In other words, one can logically infer my answer to the question, "Why be moral?" from Plato's ancient doctrine of the form of the good. But because neither Plato (nor, to the best of my knowledge, anyone

20 Nietzsche 1988, *TI* vol. 6, p. 155/X: 2; vol. 6, p. 68/II: 2; vol. 6, p. 72/II: 10.
21 Nietzsche 1988, *BGE* vol. 5, p. 17/3; cf. *GM* vol. 5, pp. 400–02/III: 24–25, *GS* vol. 3, pp. 574–77/344.

else thus far) has actually advocated this answer in responding to the question, "Why be moral?" I have immodestly billed my answer as a new answer. However, whether the answer is new or in fact extremely old is ultimately not terribly important, philosophically speaking. The important issue is whether the answer is correct. And I have tried to show that is.

Moral/Nonmoral

One key objection that always comes up when I raise these issues is the following: "But even if we grant your claim that norms and values are entailed by meaning and rational activity, how do you know the norms and values you are talking about are *moral* norms and values? Not all norms are moral norms, not all values are moral values." This is an important question, and I certainly agree that there are many different kinds of norms – e.g., aesthetic norms, legal norms, social norms, norms of etiquette, norms (rules) of tennis and every other sport. And the same holds for values. A good baseball player is not necessarily a morally good person, a morally good person is not necessarily a good philosopher, etc. And I also agree that we should not try to collapse these different kinds of norms and values into one unwieldy blob. To do so constitutes a distortion of reality. These claims about the plurality of norms and values seem intuitively true and beyond dispute.

However, despite the intuitive obviousness of these claims, the trickier question of how exactly to distinguish moral norms and values from nonmoral norms and values is not so easy to answer, in part because the concept *moral* itself is essentially contested (Gallie 1956). For instance, in *The Definition of Morality*, a collection of late twentieth-century essays devoted to defining morality, editors G. Wallace and A.D.M. Walker conclude their introductory remarks about the contributions by noting:

> Our review of the various types of definition [of morality] completed, it can be seen that none of the theses we have examined states a sufficient condition for a rule or principle's being a moral rule or principle, and that understood as stating a necessary condition, many of them present difficulties. (Wallace and Walker 1970, p. 20)

Keeping this skeptical observation about the essential contestability of the concept of morality in mind, let's see if we can nevertheless make some headway in answering the objection that the norms and values we have referred to are not necessarily *moral* norms and values, using Wallace and Walker's collection of different definitions of morality as a benchmark (see Wallace and Walker 1970,

pp. 8–19). Let's start by examining the kind of ought that is entailed by meaning claims. Hill writes:

> What is the nature of the 'ought' that Gibbard relies on in explaining ... [his] normativity thesis? He tells us that it is exceptionless, in the sense that statements involving it hold across all contexts and for all agents. It isn't an ought that applies only to people who are interested in believing the truth, or only to people who want to have degrees of conviction that are in line with the weight of evidence. It isn't an ought that applies only when agents have certain interests or occupy a certain perspective. Moreover, it is an 'all things considered' ought ... (Hill 2013, p. 2; cf. Gibbard 2012, pp. 12–16)

An exceptionless ought that holds across all contexts and for all agents? An all things considered ought? We're talking about a very strong ought here, and according to many traditional conceptions of 'moral', it sure sounds a lot like a moral ought. For instance, this kind of ought is universalizable – it applies to everyone in relevantly similar situations (see Wallace and Walker 1970, pp. 8–9). And, since we're talking about an *ought*, it is of course also prescriptive – it tells us what we ought to believe and how we ought to act (see Wallace and Walker 1970, pp. 9–10). But does this kind of ought override other oughts (see Wallace and Walker 1970, pp. 10–11)? It would seem to – e.g., the imperative to disbelieve contradictions (if you believe snow is white, you ought not to believe that nothing is white) would seem to override an ought of etiquette such as "Which fork should I eat the salad with?" For if we fail to adhere to the former ought, we won't even be able to ascertain correctly which cases involves violations of the second ought. Does this kind of ought have an importance that other kinds of ought lack (see Wallace and Walker 1970, pp. 11–13)? Again, it would seem to (e.g., disbelieving contradictions seems intuitively more important than using the correct fork at dinner). But are there certain types of sanction associated with this ought that are not associated with other kinds of ought (see Wallace and Walker 1970, pp. 14–16)? For instance, social pressure in the form of ostracism or hostility, or personal feelings of guilt, shame, and remorse? In many cases, yes. Think of the public outcry, legal repercussions, and personal shame that often accompany scientists who distort data, or politicians who lie out of self-interest, or even students who cheat on exams. To cite just one example, recall the trial of Eric Poehlman, reported in the *New York Times Magazine*, a tenured science professor at the University of Vermont who "pleaded guilty to lying on a federal grant application and admitted to fabricating more than a decade's worth of scientific data on obesity." Sentenced to "one year and one day in federal prison, followed by two years of probation," the Poehlman episode grew into "one of the most expansive cases of scientific fraud in U.S. history" (Interlandi 2006, p. 1). And does this ought refer to a specific kind of content that is

absent in the case of other oughts, e. g., the promotion of social harmony (see Wallace and Walker 1970, pp. 16 – 19)? Again, it would seem to. Adherence to the norms of truth and respect for evidence does promote social harmony – indeed, as Hobbes reminds us, without this adherence and respect civil society becomes impossible.

However, because Wallace and Walker's collection is limited to competing late twentieth-century definitions of morality, it does not include that old Kantian standby, the categorical imperative. Are the oughts of meaning and rationality "objectively necessary" by themselves (Kant 1996a, p. 67/4: 414), without reference to any other ends? Will they "set human beings in motion irrespective of other desires or obligations they may have," and do they "uniformly trump all other considerations" (Glock 2005, p. 240)? This is an extremely high hurdle. Although we have shown both that those individuals who completely disavow the norms embedded in meaning and rationality cannot even coherently express their disavowal without subscribing to the very norms from which they are trying to escape, and that human society itself is not ultimately viable without adherence to these norms, neither of these truths can establish the stronger claim that the normativity of meaning and rationality is a categorical imperative. These oughts are not objectively necessary by themselves. For instance, those who do not want to be successful in interpreting the world or in communicating coherently with others (or even with themselves) need not consistently abide by these norms and values.[22]

So where does this leave us? According to many contemporary definitions of 'moral', the norms involved in meaning and ordinary cognitive activity do count as moral. Only when judged by the sterner test of Kant's categorical imperative do they fail. But it is no secret that this particular yardstick has failed to win universal support, and some contemporary philosophers maintain that those who hold that moral oughts are objectively necessary by themselves are merely "relying on an illusion" in an attempt to give these oughts "a magic force" (Foot 1978, p. 167). So perhaps a less controversial test is needed.

In showing that arguably moral norms and values are involved in ordinary cognitive activities and meaning claims – activities and claims that are often viewed as paradigmatically nonmoral – we have also succeeded in finding a new and different answer to the age-old question, "Why be moral?" If this argument is accepted, it does show that people have compelling reasons to try to be moral in at least a minimal sense – not merely because it is in their self-interest to do so (though it often is), or because they will be happier for doing so (though

22 I would like to thank Beatrix Himmelmann for pressing this point on me.

they often will be), or because it is morally right to do so (though it is) – but for other fundamental reasons. People who forswear moral norms and values completely will ultimately not be able communicate coherently with others or utter meaningful statements, and they will thus be cut off from a great many activities that are central to human life.

But there is of course a price to be paid for this way of answering the question. In arguing that the scope of morality is in fact much broader than commonly assumed, we will be accused of an unwelcome moralism – "seeing things as moral issues that aren't, and thereby overmoralizing the universe" (Coady 2008, p. 17), or "thinking about morality in ways that discount the importance of other (non-moral) values" (Taylor 2012, p. 2; cf. 4, 127–29). And if the implicit assumption concerning the limited scope of morality and the accompanying definition of 'moralism' as a denial of this assumption are both accepted, then a conviction on this charge seems inevitable. However, I reject both the assumption and the definition. For the assumption is question-begging, and the definition does not adhere to ordinary usage. And once a more traditional definition of 'moralism' is substituted – viz., the inordinate desire to make judgments about others' morality, particularly in cases where common sense deems it inappropriate to do so – then an acquittal is quite probable. For nothing said in this essay even involves making moral judgments about the conduct of others, much less doing so in inappropriate circumstances. There are important conceptual and normative arguments for steering clear from this sort of traditional moralism, ranging from Kant's repeated warnings about the inscrutability of our moral status and the opacity of human intentions,[23] to Jesus' insightful question: "Why do you look at the speck of sawdust in your brother's eye, with never a thought for the great plank in your own?" (Matthew 7: 3–4).[24] Rather, my concern has been to show that there are fundamental moral norms and values that are presupposed by many mundane human activities, and that much as we might prefer to avoid "messy morality" (Coady 2008), in real life this is very, very difficult to do.

Also, in arguing that moral norms and values underlie ordinary cognitive practices, we are not guilty of collapsing all types of norms and values into one big unwieldy blob. The norms of etiquette remain distinct from legal norms, the rules of tennis are not the same as the rules of chess, aesthetic values are not the same as moral values, etc. And the question of which of these norms or values should override the other or be deemed more important in cases of

23 Kant 1996a, p. 61/4: 407; p. 593/6: 392; Kant 1996b, p. 95/6: 51; cf. O'Neill 1989, pp. 7, 85, 88, 130.

24 In *The New English Bible with the Apocrypha*. 1970. New York: Oxford University Press and Cambridge University Press.

conflict still remains an open question, at least according to the modest position defended in the present essay.

Finally, for those who believe that there is more to life than morality, and that the nonmoral goods available to humans are ultimately more significant than the moral goods (Wolf 1982),[25] the thesis defended in this paper is not a direct refutation of their views. I have merely tried to show that moral norms and values need to be adhered to even in the pursuit of nonmoral goods, to the extent that such pursuit involves making meaningful claims and communicating rationally. The question of which goods are ultimately more significant is a separate issue.[26]

Bibliography

Annas, Julia. 1981. *An Introduction to Plato's Republic*. Oxford: Clarendon Press.

Aristotle. 1999. In Terence Irwin (trans. and ed.) *Nicomachean Ethics*, 2nd ed. Indianapolis: Hackett.

Baier, Kurt. 1958. *The Moral Point of View*. Ithaca, NY: Cornell University Press.

Blackburn, Simon. 1984. *Spreading the Word: Groundings in the Philosophy of Language*, Oxford: Clarendon Press.

Boghossian, Paul A. 2005. "Is Meaning Normative?" In Christian Nimtz and Ansgar Beckermann (eds.), *Philosophie und/als Wissenschaft*, 205–218. Paderborn: Mentis.

Bradley, F. H. 1927. *Ethical Studies*. 2nd ed. London: Oxford University Press.

Camus, Albert. 1955. *The Myth of Sisyphus and Other Essays*, trans. J. O'Brien. New York: Alfred A. Knopf.

Coady, C. A. J. 2008. *Messy Morality: The Challenge of Politics*. Oxford: Oxford University Press.

Fielding, Henry. 1950. *The History of Tom Jones, a Foundling*. New York: Modern Library.

Foot, Philippa. 1978. *Virtues and Vices and Other Essays in Moral Philosophy*. Berkeley and Los Angeles: University of California Press.

Frankena, William K. 1980. *Thinking about Morality*. Ann Arbor: University of Michigan Press.

Gallie, Walter Bryce. 1956. "Essentially Contested Concepts." In *Proceedings of the Aristotelian Society* 56, 167–198.

Gauthier, David P. (ed.) 1970. *Morality and Self-Interest*. Englewood Cliffs, NJ: Prentice- Hall

25 Wolf's essay, "Moral Saints," is also reprinted as the last selection in the "Why Be Moral?" section of Shafer-Landau. The editor interprets her as encouraging "us not to be" moral (Shafer-Landau 2007, p. 145).

26 Earlier versions of this paper were presented as invited lectures at the Why Be Moral? conference held at the University of Tromsø in September 2013 and at the University of New England in February 2015. I would like to thank Beatrix Himmelmann and Andrew Rotondo for their detailed comments on my written text, as well as audience members at each venue for their valuable questions following my presentations.

Gert, Bernard. 1988. *Morality: A New Justification of the Moral Rules*. New York: Oxford University Press.
Gewirth, Alan. 1978. *Reason and Morality*. Chicago: University of Chicago Press.
Gibbard, Allan. 2012. *Meaning and Normativity*. New York: Oxford University Press.
Glock, Hans-Johann. 2005. "The Normativity of Meaning Made Simple." In Christian Nimtz and Ansgar Beckermann (eds.), *Philosophie und/als Wissenschaft*, 219 – 241. Paderborn: Mentis.
Habermas, Jürgen. 1991. *Erläuterungen zur Diskursethik*. Frankfurt: Suhrkamp.
Hart, H. L. A. 1961. *The Concept of Law*. Oxford: Clarendon Press.
Hill, Christopher S. 2013. Review of Allan Gibbard, *Meaning and Normativity*. In *Notre Dame Philosophical Reviews* 2013.07.20. http://ndpr.nd.edu/news/41184-meaning-and-normativity/.
Hills, Alison. 2010. *The Beloved Self: Morality and the Challenge from Egoism*. Oxford: Oxford University Press.
Hobbes, Thomas. 1946. *Leviathan*, ed. by Michael Oakeshott. Oxford: Basil Blackwell.
Hospers, John. 1970. "Why Be Moral?" In Hospers and Sellars (1970), 730 – 746.
Hospers, John. 1961. *Human Nature: An Introduction to the Problems of Ethics*. New York: Harcourt, Brace, & World.
Hospers, John and Wilfrid Sellars (eds.). 1970. *Readings in Ethical Theory*. 2nd ed. New York: Appleton-Century-Crofts.
Interlandi, Jeneen. 2006. "An Unwelcome Discovery." http://www.nytimes.com/2006/10/22/magazine/22sciencefraud.html?_r=0 , visited on 20 July, 2013.
Kant, Immanuel. 1997. In Peter Heath and J. B. Schneewind (eds.), *Lectures on Ethics*. New York: Cambridge University Press.
Kant, Immanuel. 1996a. In Mary J. Gregor (trans. and ed.), *Practical Philosophy*. New York: Cambridge University Press.
Kant, Immanuel. 1996b. In Allen W. Wood and George di Giovanni (trans. and eds.), *Religion and Rational Theology*. New York: Cambridge University Press.
Kavka, Gregory. 1984. "A Reconciliation Project." In David Copp and David Zimmerman (eds.), *Morality, Reason, and Truth: New Essays on the Foundations of Ethics*, 297 – 319. Totowa, NJ: Rowman & Allanheld.
Korsgaard, Christine M. 2006. "Morality and the Distinctiveness of Human Action." In Frans de Waal, *Primates and Philosophers: How Morality Evolved*, eds. Stephen Macedo and Josiah Ober, 98 – 119. Princeton: Princeton University Press.
Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
Lance, Mark Norris and John O'Leary-Hawthorne. 1997. *The Grammar of Meaning: Normativity and Semantic Discourse*. New York: Cambridge University Press.
Louden, Robert B. 2013. "Meaningful but Immoral Lives?" In Beatrix Himmelmann (ed.), *On Meaning in Life*, 23 – 43. Berlin: de Gruyter.
Louden, Robert B. 2012. "Are Moral Considerations Underlying?" In Sebastian Schleidgen (ed.), *Should We Always Act Morally? Essays on Overridingness*, 117 – 137. Marburg: Tectum Verlag.
Louden, Robert B. 2011. *Kant's Human Being: Essays on His Theory of Human Nature*. New York: Oxford University Press.

Louden, Robert B. 1992. *Morality and Moral Theory: A Reappraisal and Reaffirmation*. New York: Oxford University Press.
McDowell, John. 1984. "Wittgenstein on Following a Rule." In *Synthese* 58, 325 – 364.
Nelson, Stephanie A. 1998. *God and the Land: The Metaphysics of Farming in Hesiod and Virgil*. With a Translation of Hesiod's *Works and Days* by David Grene. New York: Oxford University Press.
Nielsen, Kai. 1970. "Why Should I Be Moral?" In Hospers and Sellars (1970), 747 – 768.
Nielsen, Kai. 1958. "Is 'Why Should I Be Moral?' an Absurdity?" In *Australasian Journal of Philosophy* 36, 25 – 32.
Nietzsche, Friedrich. 1988. In Giorgio Colli and Mazzino Monitnari (eds.), *Kritische Studienausgabe* (= KSA), 15 vols. Berlin: de Gruyter.
Ogden, C. K. and I. A. Richards. 1923. *The Meaning of Meaning*. London: Kegan Paul.
O'Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
Pahel, Kenneth and Marvin Schiller (eds.). 1970. *Readings in Contemporary Ethical Theory*. Englewood Cliffs, NJ: Prentice-Hall.
Parfit, Derek. 2011. *On What Matters*. 2 vols. Oxford: Oxford University Press.
Plato. 1996. In John Cooper (ed.), *Complete Works*. Indianapolis: Hackett.
Pojman, Louis P. 1990. *Ethics: Discovering Right and Wrong*. Belmont, CA: Wadsworth.
Pojman, Louis P. (ed.). 1995. *Ethical Theory: Classical and Contemporary Readings*. 2nd ed. Belmont, CA: Wadsworth.
Prichard, H. A. 1968. *Moral Obligation and Duty and Interest: Essays and Lectures*. Oxford: Oxford University Press.
Quinn, Philip L. 2008. "The Meaning of Life According to Christianity." In E. D. Klemke and Steven M. Cahn (eds.), *The Meaning of Life: A Reader*, 35 – 41. 3rd ed. New York: Oxford University Press.
Ricken, Friedo. 2010. *Warum moralisch sein? Beiträge zur gegenwärtigen Moralphilosophie*. Stuttgart: W. Kohlhammer.
Sellars, Wilfrid. 1956. "Empiricism and the Philosophy of Mind." In Herbert Feigl and Michael Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, I: 253 – 329. Minneapolis: University of Minnesota Press.
Shafer-Landau, Russ. 2010. *The Fundamentals of Ethics*. New York: Oxford University Press.
Shafer-Landau, Russ (ed.). 2007. *Ethical Theory: An Anthology*. Malden, MA: Blackwell.
Sher, George (ed.). 2012. *Ethics: Essential Readings in Moral Theory*. New York: Routledge.
Sidgwick, Henry. 1907. *The Methods of Ethics*. 7th ed. New York: Macmillan.
Stace, W. T. 1937. *The Concept of Morals*. New York: Macmillan.
Stephen, Leslie. 1907. *The Science of Ethics*. New York: G. P. Putnam's sons.
Taylor, Craig. 2012. *Moralism: A Study of a Vice*. Montreal: McGill-Queen's University Press.
Taylor, Paul W. 1978. *Problems of Moral Philosophy: An Introduction to Ethics*. 3rd ed. Belmont, CA: Dickenson.
Thornton, J. C. 1970. "Can the Moral Point of View Be Justified?" In Pahel and Schiller (1970), 442 – 453.
Toulmin, Stephen. 1950. *An Examination of the Place of Reason in Ethics*. Cambridge: Cambridge University Press.
Wallace, G. and A. D. M. Walker (eds.). 1970. *The Definition of Morality*. London: Methuen & Co.

Waismann, Friedrich. 1976. *Logik, Sprache, Philosophie*. Stuttgart: Reclam.
Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, trans. G. E. M. Anscombe. 3rd ed. New York: Macmillan.
Wolf, Susan. 1982. "Moral Saints." In *Journal of Philosophy* 79, 419 – 439.

Hallvard J. Fossheim

Aristotle on Virtuous Questioning of Morality

One Thought Too Many?

It is often thought that asking oneself "Why be moral?" is a sign of ethical immaturity. Asking oneself whether one should act justly, or honestly, or bravely —or perhaps, asking oneself whether one should perform an act from the appropriate ethical motivation—appears in itself to betray some sort of shortcoming. The more you are able to think seriously outside the ethical box, as it were, the more it might seem that this is not a good thing, but rather a bad thing.

One might think of Bernard Williams's famous notion of "one thought too many" in this context.[1] His comments, as I understand them, are primarily about the questionable psyche of the person who needs a detour through moral theory (in his example: the idea of impartiality) in order to decide what is the right thing to do. But something like that picture is relevant to the "Why be moral?" question too. For asking this question in full earnestness seems to require that the agent be sufficiently alienated from, or uncommitted to, an ethical stance for him to be able to step outside it and question its validity.

Asking oneself "Why be moral?" can be characterised as a shortcoming perhaps not least from a virtue ethical standpoint. For according to a common understanding of what virtue is, it means that your moral motivations have not grown into a fully-fledged ethical character. You are, as it were, still open for suggestions coming from outside the sphere of the morally responsible, which is to say that you are not good. If something is part of your character, you cannot distance yourself from it; it has become part of who you are, and not something you can take off like a shirt or a watch. So being able to ask the question implies that morality is not settled in you as your character.

The Claims

What I would like to suggest in what follows, is that this is an undue simplification of the matter. On the contrary, I would like to try to argue for two claims of

1 Cf. the final paragraphs of Williams (1981).

which the first is difficult to unite with the "one thought too many" story just outlined. And I am going to do this with Aristotle as umpire and argumentative source.

One of the claims will be that when it comes to the individual's process of betterment, the question "Why be moral?" and its family of related questions are highly useful and form an absolutely necessary prerequisite in the process of moral improvement.

The main conclusion that I will try to make a case for is that the good person too needs to be able to ask something very like the "Why be moral?" question. That is, not only in our development towards ethical goodness, but in the state of ethical goodness itself and as an integral part of it, the person will see and feel a pull from other, non-moral options and alternatives. This is something which puts the moral stance (if we can call it that) into relief, and amounts to a capability to ask something like the "Why be moral?" question.

The Question

Before proceeding, I should say something more about how I will approach the "Why be moral?" question. I take the question not as a request for arguments, in the abstract, for why any individual, or any rational individual, should follow morality. Such arguments go in and out of fashion, and can build on, say, egotistical considerations, normative ideas of consistency, or prudence. Rather, I hope it is already becoming clear that the angle I want to test has to do with what it means for agents like ourselves—and those less good than us, and those better than us—to be asking basic questions about how to act and why.

The question "Why be moral?" is ambiguous in several ways. Let me point out a couple of the ambiguities before advancing any further. One source of ambiguity concerns whether, when I ask myself why I should be moral, I am asking whether I should from an external point of view *act* morally in a given situation I am facing, or whether I should *be* moral in the sense of being motivated by the right sort of reasons. The implications of the two can differ greatly, as one might perform the same set of external motions with very different motives.

A related ambiguity concerns whether I am asking myself the question with a view to a singular setting, or with a view to broader vistas—ultimately, my life as a whole. This ambiguity is related to the previous one in that we tend to think of *acts* as singular, while *being* one way or another characterizes the person and the person's life more broadly. But the two are not identical, since it is possible to ask about one singular situation and find that one worries about one's soul, or to ask about one's personal qualities and find that one only cares about this be-

cause of how it might affect one's performance in the present singular situation. For present purposes, I would like to hold on to both sides of both ambiguities.[2]

Goodness, Character, and *Akrasia*

Let us start by briefly considering *akrasia* and her militant cousin *enkrateia*, as these seem, at least at first blush, to constitute Aristotle's main instances of the sort of psychological openness characteristic of the "Why be moral?" question. The acratic seems to be characterized by a two-sidedness where one of the thoughts or impulses overrules the notion that there is indeed a positive reply available. It bears noting that if it is anything like what the literature normally suggests, *akrasia* must be extremely widely diffused among us. The acratic on this conception cannot be a rare bird, but will be exemplified by most of us at least from time to time.

The acratic has standardly been used to establish a picture of Aristotle's good person as someone who no longer sees, or is able to consider, what the acratic sees and is affected by. However, we don't need to think of the motivational set of the good person as having silenced all other considerations than the ethically ideal ones. I suspect that part of what has made such readings possible has been a wrong-headed and overly broad conception of the acratic and encratic types. Aristotle's analysis of the acratic and the encratic types does not force us to admit a class into which most of us fall most of the time.

You may or you may not agree that the good person must be open to suggestions in a way not allowed by what we might call the "total silencing" interpretation. (I call it the "total silencing" interpretation, because it is not *entirely* clear to me where John McDowell's reading lands on the question of how the virtuous agent experiences the lure of alternative courses of action, or whether she does so at all.[3] What sort of experiential difference from the encratic's experience does his talk of "not count for him as any reason" (91) or "count for nothing" (92) imply? If silencing is substantially different from outweighing or overriding, then it is difficult not to see this interpretation as yielding a view of goodness

2 By these comments, I at the same time want to stress that I do not mean to limit the question "Why be moral?" to instances where the individual understands herself as asking, hyper-existentially, once and for all and across the board whether she would like to belong to morality or not. Most instances of people posing themselves the "Why be moral?" question clearly are not like this, and the ones that are, perhaps are not in the end as important as the person at the time feels them to be.

3 Cf., e.g., McDowell (1998), pp. 91–93.

which is not only idealized to the point of being non-existent, but—as as I'll try to illustrate in what follows—ethically impoverished as well.

Being Good and Seeing Bad

Here is one argument for the necessity of the good person in a certain sense having badness in him. In order to act well, one must be able to grasp the characters, motives, and schemes of others. No act is an island. What the act signifies, and not least, whether it succeeds, will depend on the agent having grasped what motivates the other individuals involved. Furthermore, success in action depends on the agent's ability to give some sort of prediction of how each of the others involved will see and react to what one does. Now this in turn depends on one being able to somehow align with each of them. Understanding another's point of view in a practical sense is not a matter of treating him or her as a black box, and through some algorithmic, statistical, or accidental approach calculating the actions that will follow. On the contrary, it is a matter of seeing the whole from that person's point of view. What are the goods at stake? What is the other person up to? Which alternative courses of action seem most desirable to him, and why? Et cetera. But you cannot do this without, as it were, having a taste of those goods, desires, and perspectives as the other person experiences them.

Whether we think of this quality or ability in terms of emotional intelligence or as being savvy, these other perspectives must be available as part of the horizon for the agent, if he or she is to be successful in action. How can we get to a closer determination of what sort of stance the good person must have towards this plethora of shortcomings?

Now "having a taste of" is not the same as "having a taste *for*" in a full sense of actually being ready to go in for whatever is the object in a given case. So the goods in question are not the individual's own preferred goods, simply because the individual is able to read the other person's perspective. But on the other hand: the way the good person relates to the other person's perspective must normally be more than *only* knowing *that* the latter person goes for that sort of thing. It must in fact involve a real perspective on ways of construing options that imply the felt presence of the "Why be moral?" question, in some cases even the denial that there is a positive account which could yield a reply to it.

Looking to Aristotle, he is clear that many less-than-perfect desires and acts are not to be liked, sympathized with, or even tolerated. Often, part of being good will be reacting instinctively to something—a personal characteristic, a motive, an action—as disgusting or as a provocation. So understanding is not the same as condoning. Nor is it, in such cases, being *overwhelmed* by the point

of view of the other. But the other's experience must still be somehow available and recognizable to the agent.

Do we have Aristotle on board here? On the one hand, Aristotle seems very clear that being good includes swift and merciless denunciation of certain actions. There is no reason to think that there is a crucial difference between how the good person is supposed to react instinctively to others' actions and how he is to react to their motives. An action, in Aristotelian parlance, normally *includes* the motive or motives behind it. And just as the person shows himself as being all the better for not having to think the matter through, but straightaway initiates the proper response, we should think (all else being equal) that the good person will be immediate in his emotional-cognitive responses to others and to their actions.

Is this picture compatible with the notion that the good person has direct access to bad motives and desires?

Emotions and States

A basic question can be articulated as: "What is it like to be a virtuous person?" On what I understand to be the common interpretation, the good person is someone who feels and reacts only in the ways that are identical to, or expressive of, virtue as a way of seeing the practical field. But we have seen that there are reasons to think otherwise. Not least is the consideration that in order to act well, you need to be able to grasp—or "see"—the point of view of all kinds of people.

This consideration also reminds us that in a way, what I take to be the received interpretation can be said to represent a sort of ethical poverty when it comes to the capacities of the good person. Now what is unavailable to the good person on this interpretation, is every point of view *other* than the good person's own, perfect, take on things. We might call this "the narrow view conception of virtue". Of course, it would all be worth it in the end, if this is what it takes to be good. But at the same time, this interpretation of the good person holds that goodness includes a massive impairment. If being in the state which allows you to act well also entails that you lack all perspectives other than that of goodness, can we then not properly speak of virtue as a sort of cognitive shortcoming?

I want to suggest that goodness-as-cognitive-impairment, in the sense just described, does not necessarily represent Aristotle's view. But I will admit that the received interpretation does have textual evidence working for it. In a manner very different from that of Plato, Aristotle seems to take as part of the core of his moral psychology that growing up and being brought up means taking on a

character, a more or less consistent set of states definable in terms of virtues and vices. And such a character is presented as a solid and lasting accomplishment. Once you have a character, changing it in any substantive way is usually presented as impossible, or close to impossible.

In order to see whether openness to the "Why be moral?" question and its concretizations in given situations might be construed as part of Aristotle's account of a virtuous person, it is necessary to consider more closely how Aristotle classifies states of character. Aristotle says,

> [s]ince there are three conditions arising in the soul—emotions, potentialities and states—virtue must be one of these. [...] If, then, the virtues are neither emotions nor potentialities, the remaining possibility is that they are states.[4]

So virtues, and by implication vices, are states. A state or *hexis* is something durable and determining for action. It is not like a coat of paint, which on a whim can be scraped off or painted over in some other colour. A state is more like a second nature.

Furthermore, a *hexis* represents a serious narrowing in relation to a capacity or *dunamis*. Says Aristotle, at a later point in the *Nicomachean Ethics*, "while one and the same capacity or science seems to have contrary activities, a state that is a contrary has no contrary activities".[5] The underlying logic seems to be that if nature provides us with a capacity to become good, then a *hexis* is a realization of that capacity. That is to say, one and the same capacity functions as the basis for different and opposing states. Capacities belong on a level of determination which can "hit both ways", as it were, in relation to virtue and vice. The same capacity thus stands as an explanation, or at least a partial explanation, of diametrically opposed ethical states. To the extent that a capacity for anger, for instance, is a basis for the virtue of mildness, the same capacity also stands in the same relation to the vice of hot-temperedness. Both the virtue and the vice, and all the other more or less vicious states building on a capacity for anger, can be traced back to that capacity as a natural condition for their existence. This is why capacities are said to relate to actions in the way the possessor of medical science relates to acts of healing: the doctor is the one eminently capable of both

4 *EN* II, v, 1105b19 – 21, 1106a10 – 12; Irwin's translation, modified by substituting, for the sake of consistency, "emotion" for "feeling" and "potentiality" for "capacity" throughout.
5 *EN* V, i, 1129a13 – 15.

killing and saving her patients. And correspondingly, a capacity has the structural role of underlying the corresponding virtue and vice alike.[6]

This would seem to constitute a strong argument in favour of the received interpretation. An ethical *hexis* determines the person. One *hexis* characterises person A and her relevant actions, while another and opposed *hexis* characterises person B and her relevant actions. So if your state is one of mildness, then mildness is all you feel in the relevant cases. Hot-temperedness is alien and nothing to you.

But does this really follow? I think there is reason to tread carefully here. The first claim seems unproblematically to follow: a *hexis* narrows a person's range of possible actions, and this is why it can function as *explanans* to the *explanandum* of her action.

But the second claim does not follow. The premises do not necessarily entail that a person with a given *hexis* does not have any access to other emotive reactions. Consider the following passages.

> in so far as we have emotions, we are said to be moved; but in so far as we have virtues and vices, we are said to be in some condition rather than moved.[7]

> First, then, neither virtues nor vices are emotions. For we are called excellent or base in so far as we have virtues or vices, not in so far as we have emotions [*kata ta pathê*]. We are neither praised nor blamed in so far as we have emotions; for we do not praise the angry or the frightened person, and do not blame the person who is simply angry, but only the person who is angry in a particular way [*pôs*]. But we are praised or blamed in so far as we have virtues or vices.[8]

There is more than one reason why it is important for Aristotle to distinguish sharply between *hexeis* (virtues and vices) and emotions. One of them is that without this distinction, there is no room for a theory of acquired ethical virtue of the sort he is out to defend. So the distinction is at least principled, in the sense that emotions and states are different even if it turns out that they are one and the same moral psychological entity in any given instance.

But this does not rule out the possibility that one might have a wide variety of emotional reactions to a given situation, as long as one of them is the ruling

6 Not all potentialities relate in this way to two diametrically opposed states. The motion of the elements, e. g., is unidirectional, in that the elements (such as a piece of earth) do not harbour a potentiality for developing more determinate states. The capacities for human goodness, however, do hold such possibilities.

7 *EN* II, v, 1106a4 – 6.

8 *EN* II, v, 1105b28 – 1106a2.

one and the one shaped into part of one's character: that is to say, *as long as one of them is the one that foreseeably determines action.*

To what extent does Aristotle think of the emotion as remaining somehow intact "underneath" the state? He cannot think of emotion as something which as a matter of fact exists in abstraction from what we might term the "how-ness" of virtue or vice. For any case of emotion must always present itself in some way or other. Aristotle says both that emotion as such belongs on a generally animal level, which is transcended by virtue, and that virtue is nothing but the alteration, rather than the overcoming, of emotion. So the question really is how much can be said about human nature, in a narrower sense, as something which is still a presence in the habituated individual. The following passage seems to support the idea of emotion as a constant aspect of, or basis for, our states and actions, inherent in us from nature. A virtue of character

> is concerned with emotions and actions, and these admit of excess, deficiency and an intermediate condition. We can be afraid, e.g., or be confident, or have appetites, or get angry, or feel pity, in general have pleasure and pain, both too much and too little, and in both ways not well; but [having these emotions] at the right times, about the right things, towards the right people, for the right end, and in the right way, is the intermediate and best condition, and this is proper to virtue.[9]

A virtuous state consists at least in part in being well off in relation to emotion. And a virtuous state marks, Aristotle seems to say, a middle point in a determinate range of states, the extremes of which are full-blown vices. Therefore, it makes good sense that at least some virtue-vice continuities get much of their unity from relating to some emotion or other.

Naturally, it is not necessarily only emotion which defines an ethical virtue. Aristotle also seems to place some emphasis on the notion of a sphere of action in this respect. But emotions do yield many of the virtues. Bravery relates in this way to fear.[10] Temperance relates to the basic appetites. Generosity, magnificence, magnanimity, and "the virtue concerned with small honours" can of course deal with goods which are ostensibly objects of appetite. But ultimately, all these virtues are also ways of relating to the emotions, such as anger, a form

9 *EN* II, vi, 1106b16 – 23.

10 And to confidence, as noted by, e.g., Urmson (1980), pp. 169 – 170. After Pears (1980) and in particular Hursthouse (1995), I do of course not mean to set forth the idea that each virtue is delimited by experiencing the right amount, quantitatively, of one emotion. What I wish to do is bring out two central aspects of the relationship between emotions and virtues/vices.

of the self-assertion of *thumos*.[11] In these central cases, then, we see that emotion is imperative in providing focus for the virtue-vice continuity.[12]

But if a virtue-vice continuity can be said to be based on emotion, that tells us something about the cognitive strength of such an emotion. The relation between emotion and state seems to be partly a logical one: that is, in a more or less completed or habituated individual, it is possible to "peel off" in analysis the state determination and be left with a notion of emotion which amounts to much more than only a reconstruction of some original natural state. To name but one example again, bravery is concerned with the emotion of fear, which is thus already conceived in terms of a desire to avoid danger. We must therefore think of Aristotelian emotions as always already more or less determinate. Expressed in terms of desire, then, fear will be a desire to avoid danger, irrespective of which state one has, that is, of how one is disposed towards one's fear. The emotion has sufficient cognitive strength to be determined apart from whether the human being in question is "well or badly off" with respect to it. On a more general level, then, the virtuous and the non-virtuous person may share the relevant emotion as a way of experiencing or grasping the situation.

There is support for such a reading outside the *Ethics* too. For this, presumably, is the level on which the *Rhetoric*'s definitions of emotions operate, allowing for its characterisation of fear as "a sort of pain or agitation derived from the imagination of a future destructive or painful evil".[13] This is a definition which determines the general cognitive direction of an emotion, that is, its general sort of object, without including an account of just how fine-grained or advanced the cognitive structure might be. But the point remains that, even in abstraction from the state, Aristotle's characterisation of fear allows it an important role in cognising the relevant situations. Sticking with the present example, a danger is something which threatens one's own being; this appears to imply that as an emotion, fear has to include some grasp of oneself as opposed to others

11 Slightly confusing in this context is how Aristotle seems to conflate honours with what seem to be merely material goods, but the clue for grasping the connection is probably that both money and honour classify as external goods.

12 More problematical is, e. g., general justice, which relates to all the emotions via its relating to all the other virtues considered as a whole, or even special justice, which is the part of general justice dealing specifically with honour or wealth or safety, and accordingly relates to emotions like anger, appetite, and fear. Not to mention a virtue like wit, which seems to relate more exclusively to a realm of action (dialogue or discussion) than to any particular recognisable emotion.

13 *Rhetoric* II, v, 1382a21 f; translation George A. Kennedy.

(whose danger is an object of pity rather than of fear); of the pain of injury or obliteration as opposed to the pain or discomfort of, for instance, mild hunger; and of a future state of affairs, that is, of something which is not yet present. But it does this without saying how the individual relates to that fear—well (the virtuous state) or badly (some more or less vicious state). This quietude is, I think, central to allowing the notion of emotion the importantly non-ethical, or rather pre-ethical, status it holds. For, given the way it is defined, an emotion like fear might apply equally to a logical analysis of the virtuous character of a mature individual and to the doings of a child (or perhaps even an animal).

So the state of affairs is something like this: in mature human beings, every external display of emotion is at the same time a display of some state or, in the case of natural virtue, state-like manner in which it is set, as a cognitive "tuning" of that emotion. Emotion seems to be treated at least partly as the aspect of a virtue or vice which is catered for by nature. A state proper, by contrast, is not the result only of nature as an internal principle of development, but has been formed by habituation. So any mature display of anger, for instance, is simultaneously a display of an emotion and of a state.

Thus, in certain contexts emotion can even be treated as if it existed in separation from virtue and vice. This is evident from Aristotle's advice to the public speaker. Aristotle's insight, in the *Rhetoric*, is that since the public speaker cannot fully take into account the individually developed character of each person in the audience, the capacities for emotional response, which are universally shared among all or most, form the level at which a reflective analysis of public address must be carried out. One person is more rash than another, and one more egoistic than another, so the address cannot be made to serve as a perfect vehicle for ensuring any fine-tuned aggressive response in each and every one among them. What the speaker must do, is attempt to create a desire for revenge in each by considering the more generally shared level of emotion. In this sense, the more general emotive level can still be said to be directly available in the mature, state-determined individual. And it may be Aristotle's view that the emotive perspectives of other characters will be available to the good or decent member of the audience, although she will not act on what has in this way been made emotively available to her. This means that a good person may be able to experience something like the impulses and perspectives not only of those who feel the existential force of the "Why be moral?" question, but even of those who have settled on a denial that the question has a positive answer. While the good person will not herself be unsettled by its force, its experiential recognition will enable her both to act and to speak more wisely.

Three Possible Objections

I want to very briefly respond to three possible objections concerning my interpretation of the specifics of Aristotle's theoretical framework. They concern, respectively, (1) the settled quality of character, (2) the difference between *praxis* and *technê* when it comes to their characteristics, and finally (3) an objection of a more purely normative sort.

As for the first possible objection, this would be the suggestion that, although we are "somehow co-responsible"[14] for our character, this very phrasing reminds us that character is not something we can change like a shirt. While this is surely true, it does bear noticing, however, that Aristotle in this passage is speaking only of people who are coming up short. He does not in fact say anything about any possible limitations of the good person. The responsibility in question concerns a less than good person who has done something that is no good, and this also means that the concern, in the context, is to point out that such a person could have made something else of himself. And in fact, even here the setting implies that the person in question is defending himself in a way that requires his ability to see that what he did was wrong. If you are blaming something on your own weakness or shortcoming, then you are certainly capable of seeing it *as* a shortcoming. So this passage will not save the "narrow view conception of virtue".

The second objection makes things more interesting. The claim here is that Aristotle himself implies a strict POV ("point of view") limitation for the good character too, in his definitional contrasting of character from technical knowledge. While technical knowledge and know-how enables you to swing both ways —say, using the same medical knowledge to save or to kill—ethically embedded insight is defined by opening up a space of action and reflection in only one direction. While the doctor can save *and* kill, the good person can perform *only* good deeds. So while the bad person is trapped in the direction of evil, the good person is equally trapped in the direction of good, unable to perform unjust, or gluttonous, or cowardly actions.

But this is a misrepresentation of the case. More accurately, the criticism depends on a confusion of levels. We can agree that a certain character is partly defined in terms of a certain set, or range, of *actions*—actions that are just rather than unjust, for instance. But this is not what is at issue. What is at issue, is whether the stability that a character seems to provide is provided through that character's inability to think and feel otherwise—that is, to think and feel

14 Aristotle's wording is *sunaitios pôs* (*EN* III, 5, 1114b23).

what would, had the experience been action-effective, lead to acts other than those of the good person. And assuredness about this latter contention is what the argument so far has hopefully served to undermine.

A final, more normative kind of criticism can also be raised. The claim will be that we do not want a state of affairs where goodness is that fragile. A world without solidly good people, without characters that are unambiguously virtuous and will remain so at least in the near future, is a world of distrust and anxiety. If there is no solidity in goodness, why bother disciplining oneself or providing others with a good upbringing?

Here, it is of course possible to respond simply that our not liking something is not necessarily a knock-down argument against its being true. But we can also say more. For again the objection misses the mark, much in the way the last objection did. The vision of a world off its hinges does not follow from the suggestion that if character entails a significant measure of stability, it does not do so through simple constriction. I want to end by trying to make this difference in levels more clear by taking a developmental perspective on character.

Insights from a Developmental Perspective

I hope to have established that Aristotle does not necessarily claim that the good person only has access to emotions perfectly attuned to performing good actions. And correspondingly, nothing in what he says denies that the less-than-good person might have emotive access both to material approaching the good person's effective state, and to the perspectives of other, alternative, less-than-good characters. On the contrary: in the case of less than good people, how could we grasp their betterment as taking place *without* an appreciation of other perspectives than those they act upon?

This also means that Aristotle does not necessarily hold that it is only the akratic and the enkratic who have experiential access to two or more emotive perspectives. What distinguishes the akratic and the enkratic is not that they have this *emotive access*, but that they are seriously in two minds about how to *act*.

We get support for this reading of what is crucial to virtue—and to vice—by considering something which Aristotle himself sees as a central issue in his discussion of virtues, namely, the question of how we develop them. It is striking, once you think about it, that Aristotle does not present any specifically mental exercise as paramount to developing virtue.

> A state [of character] arises from [the repetition of] similar activities. Hence we must display the right activities, since differences in these imply corresponding differences in the states.[15]
>
> Virtue of character [i.e. of *êthos*] results from habit [*ethos*]; hence its name "ethical", slightly varied from "*ethos*".[16]

Aristotle's unambiguous response to the question of what is required to become virtuous does not include any exercises designed to purify one's soul or to cleanse one's psychological set-up. What he presents as the all-important thing is ensuring that the acts one performs are uniformly expressive of virtue (partly defined in terms of wishing the act for its own sake). Not only is there nothing here which suggests that there is no room in the good person for alternate perspectives, as long as they do not seriously interrupt the flow of virtuous acts; the training method would seem to be positively ideal for *not* rooting out all plurality of perspective, as long as the predictable, harmonious solidity of good agency is ensured.

So while such solidity naturally requires feeling the right pleasures and pains, this does nothing to exclude a richer emotional life than the narrow view conception of virtue would allow. There is plenty of room in the logical space between Goody Two-Shoes and the enkratic for good agents who can still see and feel the attractions of other routes than that defined by virtuous action.

Emotional *Phantasia*

Do we have any more positive evidence that Aristotle might harbour a view like the one I have tentatively been attributing to him? This brings us to the topic of *phantasia*, a concept that has been understood by readers of Aristotle in widely different ways. It would go far beyond the present inquiry to try to construct anything like a unified Aristotelian theory of *phantasia*. But the few points directly relevant to our question might be brought out without attempting anything so ambitious. For as it turns out, what he has to say on the matter in *De Anima* (III.3, 10 – 11) and in *De Motu Animalium* (esp. 6 – 8) strongly suggests a psychological model in line with the sort of openness I have been advocating. Consider the following passages.

15 *EN* II, i, 1103b21 – 23.
16 *EN* II, i, 1103a16f.

> This, then, is the way that animals are impelled to move and act: the proximate reason for movement is desire, and this comes to be either through sense-perception or through *phantasia* and thought.[17]
>
> For the affections suitably prepare the organic parts, desire the affections, and *phantasia* the desire; and *phantasia* comes about either through thought or through sense-perception.[18]

The sketch is rudimentary, but clear. What should strike us first of all is that the action-inducing desire is *not* present at the outset. Rather, the more unwieldy and manifold power of *phantasia* generates a desire (*orexis*), which then in turn generates an emotional reaction (*pathos*) proper. This means that *phantasia* must also be seen as part of the early searching phase (*zêtêsis*) of action. In this phase, several possibilities might suggest themselves to the agent. We must imagine that this is the rule rather than the exception: most cases of practical searching will lead to more than one option, although they will not all be actively endorsed by the agent in the sense of generating what we might call the "ruling passion"—the action-efficient reaction.

A search may be instigated by the agent's own broad practical agenda, or by a physical state, or by something the setting unexpectedly presents him with. To concretize, the search can be initiated through a general notion to do some good, or as a reaction to an organic lack of water, or by a lion suddenly jumping up in front of him.

So the story does not start with a determinate, character-defining desire generating an action. This belongs only to the final stages of the story. In the beginning is the much more open-ended phenomenon of the imagination doing its job of displaying various hypothetical situations or results.[19] That it makes sense to think that the prospects come in the plural is also suggested by the basic example of Aristotle's, coming by the help of *phantasia* to see that "This is drink", and then drinking (*MA* 7, 701a32–33). Surely, in many situations there will be more than one option. And in those cases, the narrowing down takes place *after phantasia* has done the job of conjuring up the various possibilities.

The progression *phantasia* → *orexis* → *pathos* does not rule out that there is some emotive activity in the agent during the initial phase before *orexis* and then *pathos* are generated. Rather, Aristotle seems to suggest that there may be appe-

17 *De Motu Animalium* 7, 701a33–36; Nussbaum's translation.

18 *De Motu Animalium* 8, 702a17–19.

19 That what is imagined are resulting states which also include the agent, is suggested by the specification at *EN* III (10, 1118a16–26) that the lion does not enjoy the look or sound of the lamb, but "the prospect of a meal" (*hoti boran hexei*).

tites, thumetic impulses, and more intellectual desires present here too. Again, what is effected by the agent fastening upon a given *phantasia* is not that only then does anything like an emotional reaction start—but, that only then has the action-effecting desire and emotion come to rule the day.

Conclusion

If my suggestions make for a viable interpretation, then the question "Why be moral?", surprisingly, *does* have a place in the life of Aristotle's good agent. Although she will not be led into temptation by it, she will be able to ask this and similar questions with a seriousness and understanding available only to the person who can see and somehow feel their pull. And this ability, on the part of the good agent, makes a real difference both for the richness of her horizon and for her practical ability to realize the perspectives of other agents less good than herself.

Presumably, there will be a limit to what sort of questions the good agent can ask herself without it being indicative of a lack in goodness; similarly, there will be a limit to how insistently she can pose them. But these limits seem to be well outside the borders set by ethically ideal agency as this is traditionally conceived. Far from being a sign of moral immaturity, then, a questioning and self-questioning from well outside the safe centre defined by virtuous motivation can be seen as a prerequisite for full ethical agency. Giving a twist to Williams's formulation, we might say that in this context, something like "one thought too few" is in fact what poses the greater threat.[20]

Bibliography

Hursthouse, Rosalind (1995): "A False Doctrine of the Mean". *Proceedings of the Aristotelian Society* 81/1, pp. 57–72.

McDowell, John (1998): "Are Moral Requirements Hypothetical Imperatives?" In: McDowell: *Mind, Value, and Reality*. Cambridge, Massachusetts: Harvard University Press, pp. 77–94.

Pears, David (1980): "Courage as a Mean". In: A.O. Rorty (Ed.): *Essays on Aristotle's Ethics*. Berkeley: University of California Press, pp. 171–187.

20 I am grateful to the participants of the "Why Be Moral?" conference held at UiT, The Arctic University of Norway for their helpful suggestions, and to the editors for their valuable advice in the final revision of this piece.

Urmson, J.O. (1980): In: A.O. Rorty (ed.): *Essays on Aristotle's Ethics*. Berkeley: University of California Press, pp. 157–170.
Williams, Bernard (1981): "Persons, Character, and Morality". In: Williams: *Moral Luck*. Cambridge: Cambridge University Press, pp. 1–19.

Ivar Russøy Labukt

The Egoistic Answer

"Because it is in your own best interest" is arguably the oldest and most frequently proposed answer to the question "Why be moral?" Some philosophers have been perfectly content with this answer. But as the contributions to this volume illustrate, most of us would like to go beyond it. I sympathize with this approach, but I am skeptical about its likely success. Thus, I give an egoistic answer to the question "Why be moral?" not because I find it immediately attractive, but because I think it is, ultimately, the only kind of justification of morality that is available.

Defending my skepticism towards non-egoistic answers is too extensive a task for this paper. Instead, my aim will be to describe and assess the egoistic alternative. I will argue that it is more satisfactory than most philosophers seem to think, but still not quite as attractive as some of its proponents have claimed. Somewhat more precisely, I want to show that egoistic considerations in most cases support a commitment to morality that is fairly deep and at least as extensive as the one displayed by most actual people. At the same time, there are a number of notable limitations on an egoistic justification of a moral way of life: it does not apply to all persons and circumstances, and it does not cover all aspects of morality equally well. It may also have a significant metaethical cost: as we will see, if the authority of morality stems from self-interest, it becomes more difficult to defend the idea that the content of morality is objective, or has a realist status.

Determining the extent to which egoistic considerations support morality is not only important to those who share my skepticism of non-egoistic positions. Many of the philosophers who believe that morality has independent, non-egoistic authority admit that this authority is not always supreme: what we have most reason to do all-things-considered also depends on other perspectives, such as that of self-interest. Thus, in order to show that we have most reason to be moral in a given case, we might have to show that the moral option is also egoistically attractive, or at least not egoistically disastrous. Even the authors who hold that morality is overriding, and so rationally trumps any competing perspective, acknowledge that human beings are not, as a matter of psychological fact, able to get rid of their irrationally strong self-love (see e. g. Nagel 1986, pp. 195 – 204). From a practical point of view it would thus be a good thing if egoistic consideration gave significant support to morality.

Before examining the egoistic considerations in favor of morality more closely, I need to explain how I will be using the concepts of morality and self-interest.

1 Morality

There is one conception of morality that would ensure that there are always sufficient egoistic reasons to act morally. This is ethical egoism: the doctrine that morality is ultimately about promoting one's own self-interest. There are several well-known objections to ethical egoism. The most direct is the observation that the theory seems to condone actions that we would consider flagrantly immoral. Imagine, for instance, that you are bored and the best available entertainment is to set an innocent person on fire (you have always wondered what it looks like, and your curiosity outweighs your compassion). As long as we make sure there are no bad long-term consequences ethical egoism implies that you are then morally required to do so. You would act wrongly if you sit down and watch TV instead. James Rachels claims, relying on a similar but less dramatic example, that we can reject ethical egoism simply because it is "wicked" (J. Rachels 1974, sect. 2).

There are various ways in which ethical egoists have tried to argue that their theory does not in fact recommend behavior of this kind in real life cases. But even if they were to succeed, there is the worry that they would do so in the wrong way. What the theory would say, in effect, is that I should not set fire to innocent people because doing so is ultimately bad for *me*. This isn't what we typically have in mind when we judge that performing such actions is morally wrong.

In this paper, I will assume that ethical egoism is mistaken and that morality puts certain restrictions on the pursuit of self-interest. In order to give an egoistic answer to the question "Why be moral?", we must now make a distinction between the content of morality and its authority. While admitting that the content of the correct moral standard is in some sense and to some extent impartial, the egoist holds that the only justification for being concerned with this standard is that doing so is in one's self-interest.

Another way of expressing the view is to distinguish between the moral "ought" and the practical or all-things-considered "ought". While it is trivially true that we morally ought to be moral, that need not entail that we all-things-considered ought to be. According to the version of egoism we are considering, the egoistic "ought" is the only direct input into the all-things-considered "ought". What we ought to do from other perspectives, such as morality or aes-

thetics, can only influence what we all-things-considered ought to do by first influencing what we egoistically ought to do.

We could also say that, on this view, morality is not in itself a source of authoritative reasons for action. There is of course a sense in which there is always a moral reason to do what is morally required, but according to the egoist, this reason does not have any authority on its own. Only reasons based on self-interest do.

Finally, the view is sometimes expressed by saying that it is only rational to be moral when it is one's long-term interest. In line with this, the form of egoism under consideration is often called rational egoism (see e. g. Baier 1993). The notion of rationality is, unfortunately, used in a bewildering number of different senses in the philosophical literature. It is therefore important to stress that being rational, in the present context, amounts to nothing other than acting on the strongest authoritative reasons, or doing what one all-things-considered ought to do.

Peter Singer is an example of an author who defends rational egoism. As all philosophers know, he thinks that the principle of utility is the correct moral standard. What is less well known is that he does not think that it is, in itself, a source of reasons for action. In order to show that it is generally "rational to act morally", he thus finds it necessary to show that it is generally "in our long-term interest to do so" (Singer 1993, pp. 314 – 35). Several other utilitarians have expressed similar views (see Hare 1981, ch. 11; Brink 1989, ch. 3). While there are not many authors who go out of their way to defend rational egoism, it is not uncommon that the position is treated as a serious contender (Baier 1993), or at least a position that is important to refute (Parfit 1984). Still, I think that the position has been somewhat neglected, considering the amount of attention that has been bestowed on the much less plausible ethical version of egoism.

As I have formulated it, rational egoism presupposes that it makes sense to talk about morality without authoritative moral reasons. This presupposition is actually quite controversial. Some authors think it is part of the concept of morality that moral reasons have some independent authority (Smith 1994, ch. 3; Joyce 2001, ch. 2; Cuneo 2007, pp. 36 – 9), or even overriding authority (Gewirth 1978, p. 1; Darwall 2006, pp. 97 – 9). In fact, some of the participants in the debate on egoism seem to hold that theories of morality simply *are* theories of authoritative reasons: whatever we all-things-considered ought to do is also what is morally right to do (Sidgwick 1907; Kalin 1970, p. 86; S. Rachels 2009, sect. 3; Tännsjö 2010, ch. 3).

Philosophers who accept any of these conceptual claims have two options if they wish to express rational egoism. One is to revise their concepts so that a

claim to independent authority is no longer implicit in moral principles. If this claim is taken to be so essential to moral discourse that such a revision is impossible, they would instead have to resort to scare quotes or some technical term, such as "impartially justifiable principles for interpersonal conduct". For the sake of simplicity, I will disregard these terminological complications in the following and use straightforward moral terms.

Apart from my rejection of ethical egoism, I will not make any (controversial) assumptions about the content of moral principles. I want my arguments to be relevant to supporters of all the standard deontological and consequentialist moral theories. Given that these theories differ significantly in their demandingness, this means that I cannot give a precise estimate of how much of the moral way of life goes beyond egoistic justification. I will be comparing the egoistically desirable commitment to morality not to the one displayed by the morally perfect person (whatever he may be like), but to the one displayed by ordinary actual people.

2 Self-interest

The extent to which egoistic considerations support a commitment to morality obviously depends on how we conceive of self-interest. Some conceptions of self-interest are moralized, in the sense that they contain an irreducible other-regarding element. The most well-known example is the Aristotelian idea that individual well-being is partly constituted by the exercise of certain moral dispositions.

A more radical possibility is to hold that the interests of other people are somehow fused with one's own. On the rational egoist David Brink's objective list theory of welfare, "one is better off when another's welfare is enhanced, and especially when one enhances another's welfare." (Brink 1989, p. 243) Lester Hunt, in a presentation of Ayn Rand's egoism, claims that "one's values include, as a part of them, the good of certain other people." (Hunt 1999, sect. IV; see also Rand 1964, ch. 4) There is something odd about combining this thesis about self-interest with rational egoism. We end up with the position that I have an authoritative reason to promote the well-being of other people not because their well-being in itself should matter to me (it shouldn't), but because it automatically makes *me* better off. If one wants to claim that other peoples' interests are a source of authoritative reasons, it would seem more plausible and straightforward to say that they are so in themselves, and not exclusively through their influence on one's own interests.

At any rate, I will be disregarding this and other moralized conceptions of well-being. It is on a non-moralized notion of self-interest that the question "Why be moral?" is most pressing, challenging and interesting. There are of course many non-moralized theories to choose between in the philosophical literature. However, I think that these theories generally to a large extent converge on the egoistic value of the moral commitment I will be describing below. Thus, I do not find it necessary to make any controversial assumptions about the nature of non-moralized self-interest. My claims in the following sections should be compatible with virtually all prominent versions of hedonism, preferentialism and objective list theories.

3 Egoistic Reasons to Be Moral

In this section, I describe three kinds of egoistic reasons that count in favor of a commitment to morality.

3.1 The strategic commitment

Most obviously, perhaps, there are the directly instrumental or strategic reasons: following moral rules is often necessary in order to obtain non-moral benefits or avoid punishment. Even someone who is constantly motivated by self-interest and constantly deliberating egoistically will, therefore, often obey moral rules. This is the kind of justification of the authority of morality Hobbes offers to the "fool" (Hobbes 1651, pp. 101–3). D. A. Lloyd Thomas argues that it is sufficient to solve the age-old question of why one should be moral (Lloyd Thomas 1970). Most philosophers would be fairly disappointed with this solution, however. There is a large number of cases where the immoral agent either is too powerful to fear punishment or too clever to get caught. Also, this kind of commitment to morality is entirely without psychological depth: moral principles figure only as externally given factors in one's utility calculations.

Another strategic use of moral principles is to adopt them as rules of thumb or the basis of habits. Always having to calculate whether you ought to, say, pay for your groceries would be very tiresome, and you might be tempted to take undesirable risks. It is better to simply pay without giving the matter any thought. These "moral" dispositions would also lack any substantial psychological authority; they would be just like dispositions to brush one's teeth every day or to refrain from drinking coffee in the evening. In unusual circumstances, they could easily be set aside.

I do not wish to deny that the strategic egoistic reasons for fulfilling one's moral duty are quite weighty, at least in reasonably well-ordered societies. However, as we will see, there is little doubt that most people will benefit from a more robust moral commitment.

3.2 The psychological commitment

There is an element of the human moral orientation that is not directly under rational control. Let us call this a psychological commitment. It seems to have two basic components: the general capacity for sympathy and the capacity for a conscience that prescribes or forbids specific kinds of conduct.[1]

Most people have a deep tendency to feel good when people around them are happy and bad when they are not. This tendency is not a result of a philosophical belief about authoritative reasons. It is found in very small children, so perhaps it is not even learned (though it may of course be reinforced by praise and rewards). One may possess the capacity even if one is, like the present author, entirely convinced that it has no rational basis.

Because of sympathy, many actions required or encouraged by morality are actually quite pleasant. In particular, this is true for assisting or co-operating with others. (Think of helping an old, grateful lady across the street, for instance.) I suspect that we often underestimate how enjoyable these actions are. Perhaps we tend to assume that, just because someone else benefits from our actions, there must be some loss or sacrifice on our part. The distribution of happiness is not a zero-sum game, however.

I also suspect that the belief in the independent authority of morality actually leads some people to take less pleasure in helping and co-operating than they otherwise would have done. Experiments in social psychology show that people who are given external incentives to perform pleasant activities rate the activities as less desirable than those who engage in them simply for the sake of enjoyment (see e. g. Ariely 2008, ch. 4). Thus, it does not seem too far-fetched to suppose that accepting rational egoism could lead us to enjoy at least some types of pro-social behavior more than a believer in morality with independent authority would.

---

1 This is admittedly an over-simplification: there is presumably no neat and clear psychological distinction between the two capacities, and I do not want to deny that other non-cognitive psychological mechanisms are involved in causing moral behaviour. However, the simplification seems safe enough for my present purposes.

Another consequence of the capacity for sympathy is that many wrong actions are not particularly attractive. We do not generally enjoy hurting others. The non-egoistic attitude towards morality may encourage the suspicion that if it hadn't been for authoritative moral principles, there would be nothing stopping us from committing atrocities all the time. (This is reminiscent of how some religious people fear the horrible consequences of atheism.) I don't think there are many people who would derive substantial egoistic benefits from raping, maiming and killing other human beings. It seems that the most extreme and disturbing evil actions, at least in modern history, have more often been motivated by false moral beliefs than by a desire for personal well-being.

Although one cannot, at a given moment, decide how much sympathy to feel, one may of course influence the workings of the capacity over time. Perhaps it might seem as if it would be worth the trouble to diminish or even, if possible, eradicate the tendency to feel good or bad on account of how other people are doing. After all, having this tendency does prevent one from performing certain actions that might have very good egoistic consequences. Also, we should not forget that there is such a thing as sympathetic pain, and getting rid of that would of course be a significant benefit – especially if one is surrounded by a lot of misery.

However, it must be kept in mind that a capacity for sympathy has far-reaching influence on human well-being. It is not primarily valuable because of the rushes of sympathetic pleasure we experience now and then. Its general impact on interpersonal relations is much more important. Sympathy allows us to conceive of other people as friends and co-operators, rather than just annoying obstacles to the pursuit of one's own interests. Being "in unity with our fellow creatures", as Mill calls it, satisfies a deep human need (Mill 1871, ch. 3). Also, it makes others treat us better. In sum, one could say that it raises the baseline level of satisfaction derived from social interaction.

Since the quality of one's social life is an important part of well-being on any plausible account, it would take some weighty reasons indeed to justify the egoistic desirability of having little or no capacity for sympathy at all. This is not to deny that there are people who feel too much sympathy for their own good, or that there are extreme circumstances where it would be preferable, as judged by egoist standards, to be completely oblivious to the fate of other human beings. However, I do want to deny that it is generally true that people will become happier by blunting their sympathetic sensibilities. It seems much more likely

that rational egoism would recommend a change in the opposite direction, at least in the Western part of the world.[2]

More specific dispositions to perform or refrain from performing certain classes of actions, such as theft, promise-keeping or killing, constitute another important source of moral behavior. These dispositions are not just habits: they have a much stronger hold over us. We do not, for instance, keep promises simply because this is what we have always done, but also because we feel in some way compelled to do it, or because the idea of breaking them fills us with disgust or anxiety. I will not discuss here exactly how the human conscience works. The important point for my purposes is, once more, that is not under direct rational control: coming to believe that it fails to track authoritative reasons will not automatically silence it. This means that a rational egoist also has access to its benefits. As Sidgwick expresses it, there is a distinction between

> the general impulse to do what we believe is reasonable, and special sentiments of liking or aversion for special kinds of conduct, independent of their reasonableness. [...T]here is every reason to believe that most men, however firmly they might adopt the principles of Egoistic Hedonism, would still feel sentiments prompting to the performance for social duty, as commonly recognized in their society, independently of any conclusion that the actions prompted by such sentiments were reasonable [...]. (Sidgwick 1907, 173)

This kind of commitment to morality is similar to a phobia. A person standing at the edge of a cliff may experience intense fear without believing that he is particularly likely to fall, in the same way that I still feel bad for breaking my promises even though I am completely convinced that I have no non-derivative reason to keep them. Perhaps the moral commitment, too, can be extinguished through cognitive-behavioral therapy. There are some pretty good egoistic reasons for not doing so.[3] First, a conscience keeps its owner out of a lot of trouble. Many immoral actions promise instant gratification, and habits and rules of thumbs are often insufficient to prevent us from taking undesirable risks or imprudently

---

2 Sidgwick reached the same conclusion: "enlightened self-interest would direct most men to foster and develop their sympathetic susceptibilities to a greater extent than is now commonly attained. [...I]t seems scarcely extravagant to say that, amid all the profuse waste of the means of happiness which men commit, there is no imprudence more flagrant than that of Selfishness in the ordinary sense of the term,–– that excessive concentration of attention on the individual's own happiness which renders it impossible for him to feel any strong interest in the pleasures and pains of others." (Sidgwick 1907, p. 501)

3 Utilitarians also think it is desirable to have aversions against certain kinds of conduct, such against lying, even if there is nothing intrinsically immoral about lying according to the principle of utility (see e. g. Sidgwick 1907, book IV; Hare 1981, ch. 2 – 3).

giving in to temptation. Psychopaths, who lack these firm moral dispositions but are perfectly capable of calculation, frequently end up in jail, broke or dead. Second, I think that having a conscience also provides benefits in social interaction. It partly constitutes and partly encourages a sentiment of *respect* towards other human beings. Respecting a person is not the same as caring for his well-being; it is more like conceiving him as an independent source of claims or side constraints on what you may do. This sentiment makes interpersonal relationships deeper and ultimately more rewarding.

I suspect that it is more common that human beings have too much commitment to specific moral rules than too much sympathy, as seen from the point of view of rational egoism. For instance, some people are too concerned with telling the truth for their own good. Perhaps rational egoism would, in general, recommend a somewhat looser psychological commitment to specific moral injunctions. I do not think the changes would be very significant, though.[4]

My claim that hedonistic egoism would have relatively conservative implications for our psychological commitment to morality is strengthened by the fact that we do not get to design our own psychological commitment from scratch. Part of it is presumably attributable to our genetic make-up. It is heavily shaped by parents, teachers, peers and the general culture during childhood and adolescence. This means that our character is formed more in accordance with moral standards than egoistic.[5] (Even if the people around you should be rational egoists, this does not mean that they will provide you with the set of dispositions and sentiments that best promotes your interests. Rather, they would encourage the psychological commitment to morality that is in *their* best interest, which is likely to be more substantial.) By the time one is mature enough to reflect on one's own dispositions, they are so deeply engrained that it will take quite a lot of effort to change them. For this reason, it may be a good idea not to try to correct one's psychological moral commitment, even if it should be somewhat stronger than what one ideally would want.

4 Jan Österberg also points out that the egoist must "use a set of derived normative rules to guide his daily life" and thinks that "these rules – which, to some extent, must be internalized – largely coincide with common-sense morality" (Österberg 1988, p. 103).

5 According to R. M. Hare, there is actually no difference between the moral dispositions you would want to instill in children in order to ensure that they are likely to have a happy life, and the dispositions that the principle of utility would recommend (Hare 1981, pp. 194–8). This is surely an exaggeration, but the difference may not be very great.

3.3 The deliberative commitment

Most people also display what we may call a deliberative commitment to morality: they reason about what to do and make explicit moral choices. This commitment goes beyond the psychological commitment in two ways. First, it typically aims for the actions that are in some sense really right, and not just considered to be right by one's parents or popular opinion. Our sympathy and conscience, on the other hand, can serve their egoistic purpose without tracking any kind of moral truth. A Nazi may get pleasure from being in unity with his fellow Nazis, and in certain societies it would be egoistically desirable to have a conscience that demands immoral actions like honor killing and forbids morally innocuous things like dancing or homosexual relationships. Second, if you have a deliberative commitment to morality, your moral beliefs will sometimes play some role in bringing about morally right actions. If you, say, give money to charity, then you will to a certain extent be moved by the belief that doing so is morally desirable, and not only by a sympathetic impulse, a bad conscience or a desire to develop an egoistically useful helping disposition.

Having a deliberative commitment to morality provides egoistic benefits. First, other people will tend to like you better. (It is of course true that you could achieve the same effect just by pretending to have the commitment, but doing so would be tiresome and risky.) Second, doing what one, on a cognitive level, takes to be morally right is an attractive project. It is a comprehensive and lasting way of providing structure to one's life, and it is easily shared with other people. Both these features contribute to its egoistic value. In addition, there is something particularly appealing about the content of this project. We like the idea that we could, during our short stay here, make the world a slightly better place (see e. g. Singer 1993, pp. 332–5). Or, as David Gauthier puts it, we enjoy cooperating with others on fair terms, even if our ultimate goal may be non-cooperative in nature (Gauthier 1986, pp. 330–9). This is not just because others will show their gratitude or reciprocate; they may not always do so. The mere knowledge that we have treated other people in ways that they could not reasonably reject is a source of satisfaction independently of what these other people actually say and do.[6]

One might doubt whether such a deliberative commitment to morality – even though it might be egoistically desirable – would be psychologically sustainable for a believer in rational egoism. If you are convinced that morality

6 Thomas Scanlon puts great emphasis on this value in his account of the authority of morality (Scanlon 1998, ch. 4). He does not construe it in egoistic terms, though.

does not matter for its own sake, how could you be motivated by it and find satisfaction in pursuing it?

Admittedly, the human mind *could* have worked in such a way that we only found projects attractive and enjoyable when we believed they were directly supported by authoritative reasons. But it is quite clear that it does not actually work this way. A good example is the interest many people take in sports. For instance, I sometimes play football. I know that, from a philosophical point of view, the outcomes of football matches do not really matter for their own sake. Yet this conviction does not stop me from forming a fairly strong non-instrumental desire that my team should win. I would even, in the heat of the moment, be willing to make a (moderate) net egoistic sacrifice in order to secure a victory. As a result, I get, in the long run, much more enjoyment out of football than I would with a purely egoistic motivation.

It is important to note that the non-egoistic desire is not a result of any kind of self-deception. It is not as if I say to myself: "From a philosophical point of view, it doesn't matter at all who wins the game. But if I manage to delude myself into thinking that it actually does, I might get quite a lot of pleasure out of playing. So I will temporarily adopt a false theory of authoritative reasons." This process would never work. (If it did, any project whatsoever could be a source of satisfaction.) I do not revise my normative philosophical beliefs at all; if someone were to ask me seriously, I would admit that the victory is, in itself, completely worthless. I just find myself pursuing it even so.

There are many things that have the capacity to attract human beings in this way. As John Stuart Mill observes, even money – the paradigm example of an instrumental good – easily becomes desired for its own sake, even to the point where this is detrimental to the pursuit of what one actually takes to have intrinsic value (Mill 1871, ch. 4). On this background, there should be nothing particularly surprising about the idea of an egoistic deliberative commitment to morality.[7] In fact, such a commitment should be particularly unsurprising. Whereas goals in football or the amount of money in bank accounts do not matter at all, philosophically speaking, taking into account the interest of others is not meaningless in the same strong sense to a rational egoist. After all, other

7 Though this point was convincingly made by Butler (Butler 1725, pp. 30–6) and has been repeated many times, it is still often overlooked. For instance, C. H. Whitely claims that for the egoist a commitment to morality cannot be "a self-subsistent, independent interest like his interests in chrysanthemums or golf or the love of a good woman" (Whiteley 1976, p. 96). He never explains why the commitment to morality must be more superficial than an interest in pretty flowers.

people's interests really do matter *to them*. This makes the project of treating them rightly seem much more attractive.

I do not, of course, want to suggest that rational egoism justifies a deliberative commitment to morality as strong as the one required by non-egoistic positions. Also, it is probably true that those who believe that morality is overriding, or at least has considerable independent authority, would derive a somewhat deeper and more permanent satisfaction from their moral commitment than believers in rational egoism could do. But at the risk of sounding repetitive, I do not think that the level of deliberative commitment to morality recommended by rational egoism would be very different from the one displayed by most actual people.

Some authors seem to think that this fairly moderate, on-and-off attachment to moral principles is not feasible: in order to maximize the benefits from being moral, one must have a much stronger deliberative commitment. Gregory S. Kavka, for instance, apparently holds that a rational egoist has to choose between being completely amoral and adopting "the moral way of life" to such an extent that he would consciously decide to sacrifice his life even if he could go on to experience more happiness than unhappiness (Kavka 1985, sect. II-III). David Gauthier's ideal egoist, the "constrained maximizer", has a similarly rigid commitment to moral fairness. He makes "a choice about how to make further choices; he chooses, on utility-maximizing grounds, not to make further choices on those grounds" (Gauthier 1986, p. 158). A constrained maximizer will sometimes, if he misjudges other agents' willingness to cooperate, find himself in situations where his disposition yields suboptimal results (Gauthier 1986, p. 169). He will not just be making short-term sacrifices that will be compensated by long-term gains such as attaining a desirable character or being trusted by others. The constrained maximizer will be knowingly choosing what is, on the whole, worse for him.

It is questionable whether a rational egoist's commitment to morality *could* be this strong. Even if it would be desirable in certain cases involving cooperation, we do not have the ability to force ourselves to make certain choices in the future simply by making a resolution. No matter how much I try to constrain myself today, I will not be psychologically compelled to give up a net egoistic benefit tomorrow. As Sidgwick says, when morality demands genuine sacrifices of self-interest, any "sane person, who still regards his own interest as the reasonable ultimate end of his actions" must be able to "deliberate afresh, and to act (as far as the control of his will extends) without reference to his past actions." (Sidgwick 1907, p. 174) Unless he possesses some remarkable capacity for self-deception, a believer in rational egoism could not consciously decide that he ought

to give up a life worth living (though he may end up doing so on impulse, or as a result of weakness of will, of course).

Also, it is difficult to see why such a strong commitment to morality would be egoistically desirable, as compared to the more modest version described above. In real life cases, it is very rare that we only have access to cooperation and support from other people if we have, or appear to be having, a completely rigid disposition to follow moral principles. Usually, being generally decent suffices. On a more personal level, a project need not be thought of as supremely important in order to yield satisfaction. We are clearly able to derive enjoyment from serious non-moral projects (say, having a good career) without letting the projects trump any opposing consideration. The behavior of millions of actual people seems to show that we can do the same when it comes to morality.

4 Exceptions and Limitations

So far I have discussed the egoistic desirability of a moral commitment in very general terms. On a more specific level, we find several exceptions and limitations. In this section, I review some of the most important.

4.1 Unusual agents

The egoistic reasons for being (reasonably) moral do not apply to all sorts of agents. There are people, most notably psychopaths, who seem to lack the capacity to develop a psychological commitment to morality. This is, generally speaking, bad for them, but given that they do lack the capacity, they have no egoistic reason to try to be more moral. Perhaps there is an even larger class of people who do not find morality attractive as a deliberative project. If you are one of these people, you miss out on a potentially valuable source of satisfaction and sense of purpose. However, you are not, according to rational egoism, failing to recognize the authoritative reasons for action that you really have. What happens, rather, is that your constitution prevents you from having the reasons in the first place.

4.2 Extreme situations

For people who live under extreme poverty or other similarly difficult circumstances, the benefits of being a moral person are smaller, and the costs are larg-

er. Such people will have few strategic reasons to be moral, their sympathy and conscience might do them more harm than good, and having an egoistically justifiable deliberative commitment to morality will seem like an unattainable luxury.

Even those of us who are more fortunate cannot be sure that we will never face some extreme situation where the egoistic benefits of being moral pale in comparison to the costs. However, as long as these situations are quite rare, this does not significantly affect the validity of my claims in section 3. It will still be strategically wise to be moral in many other situations, and as explained above, the pursuit of morality can be a rewarding life project even if it sometimes must be set aside. It is only when it comes to the psychological commitment that the possibility of experiencing extreme situations has some general significance. Since this commitment is not under direct rational control and cannot be turned off at will, it is liable to have suboptimal effects in cases where a lot of happiness is at stake. For instance, I can imagine situations where I would sacrifice my life in order to save my son – no matter how convinced I might be that doing so would be irrational from a philosophical point of view. But given that such scenarios are quite unlikely, they should not be accorded too much weight when determining the egoistically ideal level of conscience and sympathy. Having a compassionate relationship to my son makes me very happy in everyday life and is for that reason worth the risk, egoistically speaking.

4.3 Distant groups

The discussion so far may give the impression that there is some general level of commitment to treating others rightly that is egoistically desirable for a given agent under given circumstances. This is actually an oversimplification. Egoistically speaking, it is more important to act morally towards those who are close to yourself.

In order to explore this issue, we need to distinguish between two different questions:

(1) Against whom does one have moral duties, and what are their relative strengths?
(2) Whom does one have egoistic reasons to treat morally rightly?

The answer to (1) is determined by the correct theory of morality. According to classical utilitarianism we have, in principle, equally strong duties towards all creatures capable of suffering or happiness. Other theories are more selective; they hold that our duties to certain people are stronger than others, and perhaps

also that there are people who have no moral claim on us at all (because, say, they have violated our rights). Rational egoism has no direct implications for this debate, and I will not take part in or assess it here. (2) is the question I am interested in. Even if I have equal moral duties towards all human beings, it may turn out that it is only in my interest to be "in unity" with some of them. Let us consider some possible limitations of this kind.

Most actual people show much greater concern for people who belong to their own social group. There are good strategic reasons for doing so, since these people have a greater influence on one's well-being and are more likely to reciprocate one's moral efforts. On the psychological level, feelings of guilt and sympathy are more easily aroused when dealing with people who are, in some respect, similar to oneself. When it comes to the deliberative commitment to morality, it would be very peculiar to make a conscious decision to be strongly concerned with the moral rights of one's own social group and completely disregard the moral rights of anyone who happens to belong to a different group. However, a more moderate favoritism does appear both feasible and reasonable. For instance, a person who has been abused as a child may want to focus his moral efforts on helping people in a similar situation, even though there are other tasks that are even more urgent from a moral point of view.

Since it is widely accepted that some notion of impartiality lies at the core of morality, it has proven very difficult to give a moral justification of our current level of partialism. For instance, Peter Singer has argued quite forcefully that the failure of rich people to help those who are very poor constitutes a serious moral flaw (see e. g. Singer 1993, ch. 8). That may be so. But as we have seen, Singer also seems to accept rational egoism. And while he may be right that it is generally contrary to self-interest and so irrational to be a morally bad person, that does not entail that this particular form of moral badness is irrational.

People can be distant in a temporal as well as social sense. Assuming that we have some moral obligations towards future generations, it is clear that the egoistic reasons to fulfil these obligations are comparatively weak. There are no strategic reasons to do so, and it is only to a limited extent that people who are not yet born trigger our conscience and sympathy. From a more deliberative point of view, however, I think that the idea of making some contribution to (or at least not obstructing) the "project of humanity" is quite attractive. I think this would suffice to provide an egoist justification of many of the measures we actually take – and perhaps even some that we don't take – in order to secure the well-being of future generations.

Peter Singer has also criticized our attitudes towards another weak group: animals. He thinks we are guilty of speciesism, which he takes to be just as morally objectionable as sexism or racism (see e. g. Singer 1993, ch. 3). Perhaps this

moral criticism is correct, too. But a rational egoist needn't justify his relative lack of concern for animals with a claim that human beings have supreme moral worth; he could simply point out that he gets more happiness from being in unity with his fellow human beings than from being nice to cows and chickens. Now I do think that rational egoism recommends having *some* moral concern for animals. Partly this is an unavoidable effect of the attitude towards human beings it recommends; if you are able to bond with people and feel sympathy for them, you will tend do the same when it comes to animals. However, I see no reason for supposing that this concern should always be as strong as the moral concern for human beings. For most of us, it is easier to isolate ourselves from the suffering of animals than the suffering of other people, and many animals do not have the ability to reciprocate our concern for them to any great extent.

4.4 Morally bad people

Though there are a few exceptions, most philosophers believe that we have moral duties also towards those who do not themselves care about their duties. According to classical utilitarianism, the happiness of a serial killer is actually, in principle, just as morally urgent to promote as that of a great philanthropist. However, one may not find it in one's interest to care very much about treating bad people rightly. To be sure, since this group is not distant, there may be good strategic reasons for doing so. We do, however, tend to feel less sympathy towards people who are willing to perform actions that are very wrong and less guilt from treating them badly. Especially from a deliberative point of view, a life project of fulfilling duties towards such people offers few, if any, prospects of egoistic satisfaction. The appeal of treating others in ways that they could not reasonably reject is to a large extent dependent on the supposition that they have some desire to do the same for us. Of course, it is notoriously difficult to determine the real moral worth of people. Nevertheless I do think that it is possible, as well as egoistically desirable, to develop a moral project that focuses more on morally good people than those who are morally bad.

5 A Threat to the Realist Status of Morality?

As we have seen, rational egoism does not make any claims about what the correct theory of morality is. But it might have implications for the sense in which a theory of morality might be correct. More precisely, it is more difficult to see how

morality could have an objective, or realist, content if it has no independent authority.[8]

Several prominent moral realists have claimed that we can in fact combine realism about moral principles with a denial of their independent authority. For instance, Russ Shafer-Landau writes:

> Do [moral] facts necessarily supply us with reasons for action? It is important to see that realism *per se* is neutral on this question. Whether moral facts invariably supply reasons for action depends not on realism alone, but very importantly on which theory of practical reason one adopts. [...] Realists who reject rationalism [the doctrine that moral reasons have independent authority] should simply insist that morality is only contingently reason-giving; moral demands supply reasons for action only when they align, for instance, with one's desires or interests. This is a perfectly consistent view. (Shafer-Landau 2003, pp. 165, 192; for similar claims, see Railton 1986, sect. V; Brink 1989, ch. 3)[9]

I agree that this outlook is not inconsistent. But rejecting moral rationalism does make moral realism less plausible, I believe. As is well known, moral philosophers have fairly different ideas of what morality is ultimately about. Is it, to name some possibilities, to make the world a better place, to act in accordance with principles that would be chosen behind a veil of ignorance, to treat people in ways they could not reasonably reject, or to follow principles that we are committed to simply in virtue of being rational agents? Once we have specified the moral question in either of these directions, it may have a determinate answer, but how do we agree on the question? If we accept moral rationalism, there is a straightforward sense in which one such proposal could be objectively correct: it could be the one that constitutes a true source of authoritative reasons for action. It might turn out, for instance, that there is an independent authoritative reason to make the world a better place, but no such reason to follow deontological principles. Consequentialism would then, arguably, be vindicated in the realist sense.

If we do not believe that morality has independent authority, this way of singling out the true nature of morality is no longer available. It is not obvious what the alternative would be. To be sure, the various proposed conceptions could seem more or less reasonable, or be more or less supported by actual moral dis-

8 A bit more precisely, realism is the doctrine that moral facts are stance-independent, i.e. that they obtain independently of whether we want them to, believe them to etc. See e.g. Shafer-Landau 2003, pp. 13–18 for a more careful statement of realism.

9 Somewhat confusingly, Shafer-Landau a few paragraphs later claims that denying moral rationalism seems to involve "a conceptual error" (Shafer-Landau 2003, pp. 192–3).

course, but what basis is there for claiming that one is correct in a stronger, realist sense?

This is not a purely metaethical problem. For some people, the appeal of morality as a life project is to some extent dependent on the supposition that there is such a thing as the true morality. If moral principles are just cultural norms or the expressions of subjective attitudes, following them might seem less meaningful and yield less satisfaction.

There may be a theoretical solution to this problem of objectivity raised by rational egoism. But a related problem would then arise: following the objectively correct moral principles might not be the most egoistically attractive other-regarding project. Assume, for instance, that some deontological account of the content of moral principles is correct in the realist sense. A person with lots of sympathy but a comparatively underdeveloped deontological conscience might then find that making the world a happier place is more rewarding than fulfilling his moral duties. Conversely, if utilitarianism is the correct moral theory, someone who feels strong deontological moral impulses but little sympathy might have more reason to treat others with deontological respect than to be morally good.

It seems, then, that giving an egoistic answer to the question "Why be moral?" is difficult to combine with the notion that there is a single true morality that should have a prominent place in the lives of all rational people.

6 Conclusion

I have claimed that rational egoism implies that most people should be roughly as moral as they are. The limitations examined in the previous two sections do not affect the truth of this claim, since people as a matter of fact already limit their moral efforts in the ways surveyed. For instance, we do neglect our duties to distant groups, and psychopaths behave at least as morally badly as rational egoism would have them to. In other words, those who have been skeptical of the doctrine because of a conviction that it would have a subversive effect on morality should revise their assessment of rational egoism.[10]

10 Of course it is one question how one ought to relate to morality according to rational egoism and another question how actual people would react if they became convinced that the theory is correct. Presumably many would overlook the considerations in favor of being a moral person I have presented here. In *that* sense, rational egoism would probably have a subversive effect on morality.

From one metaphilosophical perspective, we should expect, or perhaps even demand, that a theory of rationality implies that most people are roughly rational most of the time. We could then take my discussion as tending to confirm rational egoism. From another perspective, which I do not find unreasonable, we should, given the kind of world we live in, expect philosophers to do more than merely provide a rational justification of the status quo. However, even if we have high hopes of this kind, we could still find ourselves stuck with the egoistic answer. That depends on the merits of alternative answers, and as I have said, I will not attempt to assess them here.

However, we are now in a better position to address another question raised by my discussion: could principles that rely on self-interest for their ultimate authority count as genuinely moral principles? Some of the resistance to this idea may be attributable to the worry that one could only have a very superficial and cynical commitment to principles of this kind. This is, as we have seen, not true. If my claims in earlier sections are at least roughly right, a supporter of rational egoism would in many situations be difficult to distinguish behaviorally and psychologically from a person who thinks that morality has independent authority. He will feel impulses to help others and to engage in or abstain from certain more specific kinds of conduct, and these impulses will typically have the character of what Samuel Scheffler calls "authoritative motivation" (Scheffler 1992, ch 5): the very perception of the situation is motivationally laden. Moral actions automatically come to be represented as actions that just have to be done. On a more cognitive level, the conviction that an action would be morally right automatically counts as a weighty consideration in its favor; in everyday deliberation, moral reasons do not have to be constantly ratified by egoistic calculation.

According to David Brink, the fact that a rational egoist could have much the same motivational relation to morality as a believer in moral rationalism suffices to meet the "Kantian" objection that complying with morality on grounds of self-interest would be heteronomy (Brink 1989, p. 244). However, what Kant requires from a genuinely moral person is not that some particular drive or feeling must be present whenever he acts, but that he acts on the basis of a conviction that morality has supreme authority (see Wood 2008, pp. 25–6). And this a rational egoist cannot do, of course. Even though he might acknowledge that the needs fulfilled by morality are deep and widely shared, they are not in principle different from other needs. There is no special authority that attaches to morality and not to, say, sports or fine food.

In other words, the commitment to morality I have been describing in this paper is not all that Kant and many others have taken the rational moral commit-

ment to be. Supporters of rational egoism should admit that their theory conflicts with certain pre-theoretical expectations about morality.[11] On the other hand, the egoistic commitment is deeper and more substantial than what is often supposed. It is simplistic and misleading to say that a rational egoist cannot be a moral person since he only cares about himself. I will not discuss exactly what we should say. What is philosophically interesting is the nature and extent of the egoistically ideal commitment to impartial principles, and not which terms we use to describe it.

Bibliography

Ariely, Dan (2008): *Predictably Irrational*. New York: Harper Collins.
Baier, Kurt (1993) "Egoism". In: Singer, Peter (ed.): *A Companion to Ethics*. Oxford: Blackwell, pp. 197 – 204.
Brink, David O. (1989): *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
Butler, Joseph (1726/1973): "Upon the Love of Our Neighbour". In: Milo, Ronald D. (ed.): *Egoism and Altruism*. Belmont, California: Wadsworth Publishing Company, pp. 25 – 36.
Cuneo, Terence (2007): *The Normative Web*. Oxford: Oxford University Press.
Darwall, Stephen (2006): *The Second-Person Standpoint*. Cambridge, Massachusetts: Harvard University Press.
Gauthier, David (1986): *Morals by Agreement*. Oxford: Oxford University Press.
Gewirth, Alan (1978): *Reason and Morality*. Chicago: The University of Chicago Press.
Hare, R. M. (1981): *Moral Thinking*. Oxford: Oxford University Press.
Hobbes, Thomas (1651/1996): *Leviathan*. Cambridge: Cambridge University Press.
Hunt, Lester (1999): "Flourishing Egoism". In: *Social Philosophy and Policy* 16.
Joyce, Richard (2001): *The Myth of Morality*. Cambridge: Cambridge University Press.
Kalin, Jesse (1970): "In Defense of Egoism". In: Gauthier, David P. (ed.): *Morality and Rational Self-interest*. Englewood Cliffs, N. J.: Prentice-Hall.
Kavka, Gregory S. (1985): "The Reconciliation Project". In: Zimmermann, David/Copp, David (ed.): *Morality, Reason and Truth*. Totowa, N.J.: Rowman and Allanheld.
Lloyd Thomas, D. A. (1970): "Why Should I Be Moral?". *Philosophy* 45, pp. 128 – 139.
Mill, John Stuart (1871/2000): *Utilitarianism*. Peterborough, Letchwood, Rozelle: Broadview Press.
Nagel, Thomas (1986): *The View from Nowhere*. Oxford: Oxford University Press.
Österberg, Jan (1988): *Self and Others*. Alphen aan den Rijn: Kluwer Academic Publishers.
Parfit, Derek (1984): *Reasons and Persons*. Oxford: Oxford University Press.
Rachels, Jim (1974): "Two Arguments against Ethical Egoism". *Philosophia* 4, pp. 297 – 314.
Rachels, Stuart (2009): "On Three Alleged Theories of Rational Behavior". *Utilitas* 21, pp. 506 – 520.

11 Peter Singer makes a concession of this kind in his discussion of the egoistic justification of morality (Singer 1993, pp. 322 – 6).

Railton, Peter (1986): "Moral Realism". *The Philosophical Review* 95, pp. 163 – 207.
Rand, Ayn (1964): *The Virtue of Selfishness*. New York: Signet.
Scanlon, Thomas (1998): *What We Owe to Each Other*. Cambridge, Massachusetts: Harvard University Press.
Scheffler, Samuel (1992): *Human Morality*. Oxford: Oxford University Press.
Shafer-Landau, Russ (2003): *Moral Realism*. Oxford: Oxford University Press.
Sidgwick, Henry (1907/1981): *The Methods of Ethics, Seventh Edition*. Indianapolis, Cambridge: Hackett Publishing Company.
Singer, Peter (1993): *Practical Ethics, Second Edition*. Cambridge: Cambridge University Press.
Smith, Michael (1994): *The Moral Problem*. Oxford: Blackwell Publishing.
Tännsjö, Torbjörn (2010): *From Reasons to Norms*. Dordrecht, Heidelberg, London, New York: Springer.
Whiteley, C. H. (1976): "Morality and Egoism". *Mind* 85, pp. 90 – 96.
Wood, Allen W. (2008): *Kantian Ethics*. Cambridge: Cambridge University Press.

Beatrix Himmelmann

Why Be Moral? An Argument from the Human Condition in Response to Hobbes and Nietzsche

Given what we know about human beings and their behaviour through the ages, it has been repeatedly doubted whether morality deserves the praise it commonly receives. The gap separating what has been understood by morality and what has been called the "realities behind morality" (*die Realitäten hinter der Moral*),[1] suggesting that the very idea of morality might conceal and devalue those realities, appears to be too deep and, therefore, unbridgeable. Hence, this picture might be misleading.

In contrast to the difficulties of ascertaining the substantiveness and the features of morality, it seems fairly safe to say that there is a general human striving not only for self-preservation but also for well-being or happiness, thus for some kind of satisfaction or completion. However this aspiration is dubbed, it appears to underlie human conduct of life most essentially. We want to get hold of what we deem useful or pleasant or valuable. But what we are able to seize will not necessarily last so that we could attain what satisfies our wishes and needs once and for all. On the contrary, circumstances change and so might our desires. They are not set in stone once and for all.

Not least because everything is in constant flux and nothing seems to be stable, another factor becomes important, presumably in keeping with the pursuit of happiness: appreciation and esteem for the possibility of doing the things we hope make us flourish. This possibility of doing the things we expect to be conducive to our well-being obviously includes relevant capabilities of agents. In short, we all value the power to bring about what we desire. Power in various guises such as skills, physical and mental, foresight, prudence etc. appears to be necessary in order to obtain and possibly maintain what we need, want, and wish for.

The employment of power seems to shape not only our dealings with all kinds of things or objects but affects our relations with fellow human beings as well. In fact, Max Weber defines the concept of power by starting out from these relations among human beings, apparently considering its usage with reference to our relations towards non-human things derivative: "Power (*Macht*) is

1 Nietzsche 1988e, p. 508 (= N 16[71]).

the probability (*Chance*) that one actor within a social relationship will be in a position to carry out his own will despite resistance, regardless of the basis on which this probability rests."[2] We see that Weber, when introducing the concept of power, comes straight to the crucial point. Power and its exertion originate from human willing and directly aim at human willing, which is always to be understood in terms of individual willing arising from an actor's will of his or her own. It is not to be expected that this multitude of individual human willing will be constantly in harmony. Others might turn out to be rivals interested in some of the same things we are – and sources of supply may be limited. Their interests might be opposed to ours, and so we stand in one other's way. Their interpretation of our conflict might differ from ours, and this can lead to serious discord. Always and to this day there has been a tendency to solve these kinds of disagreements by deploying force, be it in coarse or refined ways and on large or small scales. For the sake of flourishing or, sometimes, just to survive, recognizing and coping with power structures appears to be inevitable. And so it looks as if the pursuit of happiness is not conceivable without also acknowledging the realities of power.

Morality, on the other hand, seems to require you to do good, or maybe "the good", even though this might do severe damage to your pursuit of happiness and let you fall behind with regard to your position within that field of forces and powers in which you appear to find yourself. So the possibility of a divide between the realities of what looks like general human endeavour and a very different order, suggesting very different criteria of valuation and evaluation, opens up.

This potential gap has been reflected in philosophical thought from the beginning. Accordingly, doubts about the status, validity, the foundation, and scope of morality were bound to arise. In what follows, I'll be starting my investigation by taking up this sceptical perspective on moral requirements. This could be a promising way of proceeding precisely because choosing this angle will allow us, for the sake of argument, to avoid making extensive use of moral vocabulary in the first place. So there won't be the danger of being lavish with moral talk which might only presuppose what has to be shown and, in addition to this, might even contribute to obscuring specifically moral claims. For the employment of idle moral talk may be one of the most effective strategies in games that serve quite different purposes, purposes indifferent or even adverse

2 Weber 1980, p. 28 / Weber 1978, p. 53.

to moral concerns. Plato already highlighted this possibility,[3] and Nietzsche's critique of morality is based on his contempt for this sort of deceit and self-deceit.

Before dealing with peculiarities like these, though, we'll have to describe the "realities" of general human striving as honestly as possible and also in as detailed a way as necessary. Thomas Hobbes, whose "courageous mind" and "tremendous love of truth" have been rightly praised[4] and who still is a vivid source of inspiration for branches of contemporary moral theory,[5] may provide us with a first set of suggestions. Later, we'll look at Nietzsche's somewhat complementary approach. Nietzsche shares some general assumptions about the *conditio humana* put forward by Hobbes, but arrives at very different conclusions. Both Hobbes and Nietzsche, I will argue, fail to acknowledge the reasons why morality is indispensable. And it is illuminating to see why they fail. So finally and *ex negativo,* I am going to suggest a positive answer to the question "Why be moral?", an answer which, I hope to show, cannot be dismissed.

I

Hobbes recognizes and analyses the pursuit of happiness. But the "Felicity of this life," he argues, cannot be thought to consist "in the repose of a mind satisfied".[6] Accordingly, the idea of a *finis ultimus*, an utmost aim, does not make sense to him. In order to live, our desires must not come to an end, and our sensual activity and imagination will never be at a standstill as long as we continue to exist. Hence "felicity", or happiness, has to be conceived of as "a continuall progresse of the desire" switching "from one object to another" and going along with the desire's recurrent satisfaction, for which there is interminable hope. As a result of this disposition, we will be always aware of the future and of future desire—that is, we will always be concerned about what Hobbes calls "the assuring of a contented life". With respect to this worry human beings are united, however diverse their passions and their understanding of how to live a happy, or at least contented, life may be.

From all the conditions mentioned, Hobbes draws an unambiguous and blunt conclusion:

3 Plato, Rep. 361 a-b.
4 Nietzsche 1988a, p. 194 (= UB I, 7).
5 Cf. Dreier 2006, pp. 55 – 96.
6 Cf. here and in what follows: Hobbes 1991), Chap. XI, p.70.

> So that in the first place, I put for a generall inclination of all mankind, a perpetual and restlesse desire of Power after power, that ceaseth onely in Death. And the cause of this, is not always that a man hopes for a more intensive delight, than he has already attained to; or that he cannot be content with a moderate power: but because he cannot assure the power and means to live well, which he hath present, without the acquisition of more.[7]

Since this is supposed to be a "generall inclination" of humans, competition for riches, for honour and command will always be found, predisposing to contention, enmity, and war. The permanent threat of war causes fear in all and a desire for peace.

However, Hobbes explains that everyone is bound by a *law of nature,* which is suggested by reason, to abstain from what is "destructive of his life, or taketh away the means of preserving the same; and to omit that by which he thinketh it may be best preserved".[8] To put it in positive terms: For everyone it is always permissible to do—that is, everyone has a *natural right* to do—whatever he or she thinks may best preserve and further their life. Consequently, on the one hand, we cannot but enter into conflict and wage war, should our self-preservation be threatened, and yet we cannot but hunt for peace on the other. Furthermore, Hobbes is very clear about what the general condition of war among humans involves:

> To this warre of every man against every man, this also is consequent; that nothing can be Unjust. The notions of Right and Wrong, Justice and Injustice have there no place. [...] Force, and Fraude, are in warre the two Cardinal vertues. Justice, and Injustice are none of the Faculties neither of the Body, nor Mind. If they were, they might be in a man that were alone in the world, as well as his Senses, and Passions. They are Qualities, that relate to men in Society, not in Solitude.[9]

It is important to emphasize, as David Gauthier did, that Hobbes's idea of a "war of every man against every man" is not based on the assumption of any "innate hostility", but on hostility derived from the ever-possible conflict between our powers of self-maintenance. Accordingly, war is the consequence of "natural insecurity" and the "natural desire" to preserve oneself.[10]

We all know how Hobbes tried to solve the fundamental problem of human existence, as he saw it. It is necessary "that a man be willing, when others are so too, as farre-forth, as for Peace, and defence of himselfe [...], to lay down his

7 Hobbes 1991, Chap. XI, p. 70.
8 Hobbes 1991, Chap. XIV, p. 91.
9 Hobbes 1991, Chap. XIII, p. 90.
10 Gauthier 1969, p. 17.

right to all things; and be contented with so much liberty against other men, as he would allow other men against himselfe".[11] For Hobbes it is an important point that *all* men lay down their (natural) right; otherwise you might expose yourself to "prey" – "which no man is bound to" – rather than to dispose yourself to "peace". The mutual transferring of right is what we call a "contract".[12] What holds just or unjust, right or wrong for all cannot be determined until now. In order to secure the contract's enforcement, however, men need "a common Power to keep them all in awe".[13] Only if there is a common power set over the parties "with right and force sufficient to compell performance", the contract won't be void.[14] All other bonds Hobbes considers too weak.

Even though men naturally love and prefer liberty – i.e. "the absence of externall Impediments"[15] – and dominion over others, they, nevertheless, finally accept "restraint upon themselves". The reason for this is "the foresight of their own preservation, and a more contented life", leaving behind the "miserable condition of warre". [16] So men agree "to erect a Common Power" and to confer "all their power and strength upon one Man, or upon one Assembly of men, that may reduce all their Wills, by plurality of voices, unto one Will". This one man or the assembly of men are appointed to "beare their Person".[17] Men submit their wills to this one will and declare:

> *I Authorize and give up my Right of Governing my selfe, to this Man, or to this Assembly of men, on this condition, that thou give up thy Right to him, and Authorize all his Actions in like manner.* This done, the Multitude so united in one Person, is called a *Common-Wealth* [...]. This is the Generation of that great *Leviathan*, or rather [...] of that Mortall God, to which we owe under the *Immortal God*, our peace and defence.[18]

We are obligated to "simple obedience" to the sovereign, to whom the whole power of prescribing the rules of life for all individuals involved is entrusted. Liberties to do or forbear according to our own discretion that are left to us depend on "the silence" of the sovereign's law.

Does this picture – at least roughly – comply with our self-understanding? Since morality has not yet been mentioned, maybe we don't need this kind of

---

11 Cf. here and in what follows: Hobbes 1991, Chap. XIV, p. 92.
12 Hobbes 1991, Chap. XIV, p. 94.
13 Hobbes 1991, Chap. XIII, p. 88.
14 Hobbes 1991, Chap. XIV, p. 96.
15 Hobbes 1991, Chap. XIV, p. 91.
16 Hobbes 1991, Chap. XVII, p. 117.
17 Hobbes 1991, Chap. XVII, p. 120.
18 Hobbes 1991, Chap. XVII, p. 120.

self-regulation? Perhaps we could or should be satisfied with a concept of what it means to strive for a happy, or rather a contented life and the idea of right, of civil laws apt to guarantee the possibility of a peaceful, if not good life for everybody?

First, we might look at Hobbes's notion of morality. The role and place of morality within the framework of Hobbes's overall conception is not easy to ascertain. Morality is introduced in connection with the so-called *laws of nature* or "dictates of Reason", which altogether prescribe what serves our preservation and is conducive to peace.[19] And so it is fair to say that the social contract is morally required because it is only by instituting a social contract that what promotes our preservation and is conducive to peace can be realized.[20] Morality, then, functions as the foundation of right and of politics in Hobbes's understanding. His idea of morality certainly endorses his specific perception of what it is to be human, a perception radically opposed to an Aristotelian one, as has been frequently observed.[21] In Hobbes's view, man is not by nature *zôon politikon* whose flourishing has to be envisioned in terms of harmonious cooperation rather than conflict-laden confrontation and struggle for self-preservation. Thus, David Gauthier takes Hobbes to be a representative of "modern moral theory," the cast and scope of which is set by three dogmas philosophy receives from economics. According to Gauthier, Hobbes embraces all of these dogmas, and not least for this reason Gauthier considers him "the greatest of English moral philosophers".[22]

The first dogma that Hobbes indeed accepts is utility. Utility serves as a measure of subjective, individual preference. Hobbes spells it out in terms of goodness and desire or appetite:

> [W]hatsoever is the object of any mans Appetite or Desire; that is it, which he for his part calleth *Good:* And the object of his Hate, and Aversion, *Evill* [...]. For these words of Good, Evill, and Contemptible, are ever used with relation to the person that useth them: There being nothing simple and absolutely so; nor any common Rule of Good and Evill, to be taken from the nature of the objects themselves.[23]

The second dogma is the idea that rationality is about calculation and maximization: the measure of the reasonableness of an action is the extent to which it is

19 Hobbes 1991, Chap. XV, p. 111. See Chapters XIV, XV, pp. 91 – 111.
20 See also Sorell 2007, p. 145.
21 Cf. Sorell 2007, p. 146.
22 Gauthier 1990, p. 11.
23 Hobbes 1991, Chap. VI, p. 39.

conducive to the agent's ends.[24] This conception underlies Hobbes's understanding of reason and science:

> *Reason* [...] is nothing but *Reckoning* (that is, Adding and Substracting) of the Consequences of general names agreed upon, for the *marking* and *signifying* of our thoughts. [...] *Science* is the knowledge of Consequences, and dependance of one fact upon another: by which, out of what we can presently do, we know how to do something else when we will, or the like, another time: Because when we see how any thing comes about, upon what causes, and by what manner; when the like causes come into our power, we see how to make it produce the like effects. [25]

Explicitly, Hobbes introduces moral philosophy as a science understood in this way: "Morall philosophy is nothing else but the Science of what is *Good*, and *Evill*", and he adds: "in the conversation, and Society of man-kind".[26] Hobbes, as we have seen, emphasizes the *natural* subjective and individualistic valence of good and evil. When he now adds "conversation" and "society" which tie individuals together, a conventional element comes into play. Since no individual can escape being confronted with fellow human beings whose preferences may differ from and conflict with hers – "good" and "evil" may be as diverse as men are diverse, Hobbes suspects –,[27] conventional arrangements have to be in place. And we understand why, according to Hobbes's conception, the most solid structure of the social contract is morally required.

The third dogma modern moral philosophy receives from economics is the idea that interests are non-tuistic, that is, interacting persons are not conceived as "taking an interest in one another's interests," as Rawls put it.[28] Not entirely coincidentally, it is the chapter on "Power, Worth, Dignity, Honour, and Worthiness" from *Leviathan* which provides evidence for Hobbes's belief in the validity of non-tuism. The "Value" or "Worth" of man, Hobbes writes in this section, "is as of all other things, his Price; that is to say, so much as would be given for the use of his Power: and therefore is not absolute; but a thing dependent on the need and judgement of another".[29]

---

24 Cf. Gauthier 1990, pp. 11, 13.

25 Hobbes 1991, Chap. V, p. 32, 35–36.

26 Hobbes 1991, Chap. XV, p. 110.

27 Hobbes 1991, Chap. XV, p. 110–111.

28 Rawls 1999, p. 12. Cf. Gauthier 1990, p. 11. The term "non-tuistic" goes back to the economist Philip H. Wicksteed: "The specific characteristic of an economic relation is not its 'egoism,' but its 'non-tuism.'" Cf. *The Common Sense of Political Economy* (1910). London: Macmillan, Book I, Chap. 5, par. 24.

29 Hobbes 1991, Chap. X, p. 63.

Not least informed by the violent conflicts resulting from "taking an interest in one another's interests" that Hobbes had seen evolving both in England and on the continent and that had, at the time, pre-eminently religious motivations, he reduces the relationship between human beings to what he takes to be their foundation: a general concern about themselves and what fosters and feeds their self-interest the pursuit of which, as a matter of fact, unites men as much as it separates them. Following self-regarding needs, desires, and judgments requires attention to power relations according to which the "value" or "worth" of any individual is determined, as Hobbes observes. As those power relations are in flux, so will be the individual's value and worth.

It is obvious that not only incompatible tuistic interests but also conflicting ways of pursuing individual self-interest can lead to serious discord. This discord may not be brought under control by simply yielding to actual power structures, given the fact that those power structures lack solidity and steadiness. The challenge is, therefore, to develop a stable framework within which anyone's attending to his self-interest may be maximized.

Again, we arrive at the "great *Leviathan*", this strange and enigmatic political beast, which Hobbes chose to be the title figure of his most influential book. Leviathan is not only a beast,[30] though, but also a "Man", an "Artificial Man", that is a machine, a "God", a "Mortall God", the "State" as a whole, and the "Soveraigne", wielding absolute power in and on behalf of the state.[31] It may well be that Hobbes depicts his eponymous monster deliberately and calculatingly in a rather perplexing way.[32] Reason as well as fancy[33] must be employed, Hobbes suggests, when it comes to familiarizing the reader – and the citizens – with the "great *Leviathan*", since the task this creature is meant to fulfil is of utmost importance. He, and no one else, is supposed to provide the stable framework within which anyone's attending to his self-interest may be maximized. Not least for moral reasons, then, every citizen in the state must be convinced that obeying the laws he issues is tantamount to acting in their best interest, viz. their enlightened self-interest. Fear and awe, induced by the highly ambiguous image of the sovereign presented in Hobbes's *Leviathan,* should do one last thing to prompt the citizens to submit to his power. And he differs from all of them who bow down to him in that he is *not* part of the contract by which individuals, covenanting with each other, institute the state. Instead, the sovereign stands out and occupies a very special role, a peculiar position of otherness. This is "because the

30 Hobbes follows, as is well known, Job 40 – 41.

31 Cf. Hobbes 1991, Introduction, Chapters XVIII, XXVIII, pp. 9, 120, 221.

32 This is shown by Johan Tralau 2007.

33 Cf. Hobbes 1991, Chap. VIII, p. 52.

Right of bearing the Person of them all is given to him they make Soveraigne, by Covenant onely of one to another, and not of him to any of them". For this reason, "there can happen no breach of Covenant on the part of the Soveraigne; and consequently none of his Subjects, by any pretence of forfeiture, can be freed from his Subjection".[34]

The sovereign's subjects agree to their unconditional submission to his will and power in exchange for a maximum of security and stability, protecting any of them against anyone else. What they hope to gain from this move is in perfect accordance with those three dogmas that modern moral philosophy, allegedly, receives from economics and, supposedly, is deemed to accept. Submitting to the will and power of the "great *Leviathan*" will, firstly, enable any of his subjects to pursue whatever they consider "good" for themselves. They will be protected from being harassed or destroyed by others who equally seek their advantage. Submitting to the will and power of the "great *Leviathan*" will, secondly, allow for the maximization of any of his subjects' pursuit of what they deem conducive to their personal ends. As far as possible, they will be protected against any undue interference in their ambitions, since proper rational assessment provided by *science* will manage to organize individual striving in the most efficient and sophisticated way. Being neutral concerning the interests of others, however, is a precondition for the success of any such regulation which requires nothing but dispassionate calculation. Submitting to the will and power of the "great *Leviathan*" will, thirdly, enforce compliance with precisely this precondition of nontuism.

Concluding the section devoted to Hobbes's argument, we can sum up: Hobbes leaves us with an answer to the "Why be moral?" question that does not allow for any gap separating what has been called the "realities" of general human striving from the realm of morality. On the contrary, morality in Hobbes's understanding serves these realities by organizing human striving for well-being in the most efficient manner. This is to say that each individual's pursuit of well-being is protected from being blocked or annihilated by others whose preferences might be incompatible with their own. In exchange for unanimously consenting to submit their wills and their powers, apparently the sources of commonly shared destruction and misery, to the one unique will and power of the sovereign, that "*Mortall God*" Leviathan, they are all assured maximal satisfaction of their individual desires.

So, finally, it turns out that in Hobbes's account the claims of morality, right, and politics all point to the institution of a sovereign power fostering the well-

34 Hobbes 1991, Chap. XVIII, p. 122.

being of everyone by depriving everyone of his "Right of Governing [him] selfe".[35] Since well-being or happiness are supposed to be the ultimate purposes of human life, a Hobbesian style solution to the "Why be moral?" seems to recommend itself.

Among other worries two main concerns arise, nonetheless. First, could we ever conceive of ourselves as renouncing our wills, giving up on the right as well as the burden of governing ourselves? Second, is it true that we are rational animals whose most salient worry is about self-interest, caring about nothing but well-being, if not pleasure? Considering these questions will lead us to reconsider Hobbes's answer to the question, "Why be moral?" Indeed, we will eventually see that it is utterly deficient.

II

Certainly we always desire, or stronger, want this or that for whatever reasons. To want something involves having a purpose we aim to achieve. And it involves employing appropriate means in order to realize what we thus pursue. We all can and often do think about our purposes and about the means of striving for them.

Thinking about the means of pursuing our aims requires us to address technical issues. Hence it concerns theoretical problems. Some kind of obstacle might occur and disrupt our dealings with the affairs of everyday life. Or there might be an especially delicate purpose we want to achieve, calling for extraordinary attention and treatment. Those cases present us with a set of challenges. We usually meet them by thinking about what is disturbing the ordinary course of events and what measures we can take to put things straight again. Alternatively, we respond by looking into the different resources we might have to reach our goal, be it ever so difficult to pursue. It is obvious that we do not engage in this kind of thinking all the time, but only if we consider it necessary in order to cope with those smaller or larger challenges of everyday life. As mentioned before, these challenges pose technical problems and require technical responses. Reason, undoubtedly, is indispensable in order to be able to deal with them. However, it is reason in the service of the satisfaction of what we desire and the achievement of what we want that we are currently considering.

Looking back at Hobbes's project, we see that he is exclusively concerned with these kinds of technical problems, dealing with them full-scale and, accord-

35 Cf. Hobbes 1991, Chap. XVII, p. 120.

ingly, on a high level of complexity. And, indeed, his social contract theory offers a technical solution to them. It is crucial to—or, better, constitutive of—Hobbes's account that he carefully abstains from surpassing the scope of this type of reasoning. Not leaving the neutral ground of coping with non-tuistic human interests in those technical ways that their regulation involves is deemed necessary for ensuring social stability. The institution and maintenance of social stability, however, is what we are morally obliged to secure— according to this picture. Why be moral? On a Hobbesian point of view, thinking about the means of pursuing our individual aims most efficiently will lead us to accept moral conduct as one of the most powerful instruments available. Enlightened self-interest appears to be the source as well as the motivation for being moral.

It turns out, however, that we know about ourselves in ways very different from conceiving of ourselves as pleasure-maximizing rational animals. We are capable of "stepping back" from being occupied with what is "out there" in the world to be dealt with in those pleasure-maximizing technical respects. And only for this reason, because we are capable of taking a step back, are we morally evaluable and responsible animals. Even though we most often engage in just doing what we do, focused on tackling the minor or major problems that may occur, we can also address ourselves as the doers who give rise to this or that particular event "out there". We say, for instance, "*I* did it" – being proud of accomplishing something or being called to account for something caused by us. When reflecting like this, we refer to ourselves as the subject of our doings to which we thus relate. So we do not merely know about what is "out there" and – to a greater or lesser extent – how to deal with it in technical terms, but we know about ourselves as the source and the end of being concerned with it. Moreover, each of us does not only know about this with regard to himself or herself but with regard to any other such "subject" or person as well. Obviously, this sort of knowledge creates obligations – most notably in moral respects.

When we are confronted with something "out there" that we take to be of our kind – i.e. another "subject" or person also capable of being aware of herself as the source and the end of her being concerned about something –, we immediately understand that it would be inadequate to treat it as nothing but an object to be handled in these or those technical respects, serving some project or goal we are pursuing. We would anticipate its resistance to being treated like this – because it is clear to us that here we do not deal with something but with someone who is able to develop the same kind of thematising relation towards himself that is open to us. He, too, is capable of being an issue for himself and so must *demand* to be treated accordingly by anyone who knows or, better, is able to know this fact. This demand is not met if I use him as a mere instrument in my hands, or regard him as an indifferent object should he fail to be of use. It

is obvious, of course, that human beings are of use to one other in complex and sophisticated ways, and it is also true that they evade one other in equally refined fashions. And there is nothing wrong with these kinds of behaviour, at least not a priori. Nonetheless, due to the constitution of human beings there are constraints on their dealings with one other. Being able to relate to themselves by becoming an issue for themselves and thus knowing themselves, these animals cannot any longer claim a state of innocence when it comes to their doings and omissions. And precisely this knowledge goes along with moral demands.

A telling example may illustrate what is at stake. The story of Artyom, an inmate in a Russian prison camp, is desperately sad, even though apparently not untypical. He had worked as a civil engineer and fallen prey to a corruption scandal. Falsely charged with the embezzlement of a tidy sum of money that others had pocketed, he was given an eight-year sentence. His wife does not come to see him, and his children are ashamed to look him in the eye because they think he is a swindler who robbed people. Artyom yearns for the sympathy of his fellow prisoners who should best understand his grief. But they turn out to be not a bit responsive. "Come on, we all know this already, and plenty more besides! It's not exactly earth-shattering news. Your own misery is always greater, obviously, but what's that got to do with anyone else?! Anyone will lend you a hand with the everyday stuff, but as for the mental anguish – sorry, pal, you just have to learn to deal with that yourself …"[36] This is, so their argument goes, because "in this place [sc. the prison camp] every other person has exactly the same story. You simply don't have the energy or time to listen to other people's woes. […] You just don't have the time for him! And yet he doesn't seem to understand this. He goes around whining on about how hopeless he feels […]."[37]

One night it's happening. Artyom is trying to take his own life; his cellmates manage to rescue him. And now, in addition to all his suffering, he is tainted with an "attempted suicide" mark causing him even more trouble.

The conclusions the first-person narrator provides could not be more to the point:

> As for the rest of us, we avoided each other's eyes, ashamed. After all, we could have known that he was on the edge, but we chose to ignore it. Indifference is a terrible sin. It's only one short step away from the professional fish-eyed look of the unscrupulous judge who believes that the happiness of his own family is justification enough for any such "Artyoms". Can we really be at peace with ourselves, pretending that someone else's fate is no concern

36 Khodorkovsky 2014, p. 61.
37 Khodorkovsky 2014, p. 61.

> of ours? How long can a country survive when indifference becomes the norm? The time of reckoning always comes eventually.[38]

It is obvious that all three dogmas allegedly characterizing "modern moral philosophy" are completely out of touch with what is morally relevant in this case, and it's safe to assume in other cases of moral significance as well. Quite the reverse, spelling out moral goodness in terms of utility as a measure of individual preference, of maximization as the measure of the extent to which an action is conducive to the agent's ends, and of non-tuism as neutrality towards one another's interests means disregarding morality as a dimension distinguishing human relations altogether. Understanding what is crucial to human life according to merely economic patterns, accepting the idea that everything is negotiable and that there is nothing beyond arranging the flow of individually preferable goods most effectively, simply does not comply with what is constitutive of human beings. What is constitutive is not, first and foremost, the satisfaction of their various desires by means of supply with relevant goods. It is what our example conveys as "mental anguish" that makes human beings most vulnerable. The regard they have for themselves and want others to recognize, this regard rooted in knowing themselves and being an issue for themselves, is the source of their moral susceptibility. And so for these animals, for human beings, there is not only physical misery but also this specific kind of mental agony. And the latter has to be considered more significant.[39] Actually, we are well aware of this fact. We only need to consider the outstanding importance given to questions of esteem and self-esteem in human understanding and self-understanding. Rousseau is one of the philosophers devoting most careful studies to precisely this issue.[40]

Why be moral? Because knowing ourselves in the manner analysed above requires it. If each of us is able to understand that he, as well as all other human beings, are beings able to relate to themselves by thematising themselves

38 Khodorkovsky 2014, p. 62 f.

39 This is the "most essential point" Marian Turski, a survivor of Auschwitz who was invited by Germany's parliament (Deutscher Bundestag) on 26 and 27 January 2015, wanted his audience to understand. Worse than the physical pain and even worse than the killing, he said, was the degradation and humiliation of inmates, the fact that they were not treated as human beings but as if they were nothing but a "louse" or a "cockroach", considered less than ordinary animals which still might be of use, whereas bugs are of no use whatsoever but call for removal. Being treated like this, Turski emphasized, is the worst thing that can happen to a human being. (Permalink: http://dbtg.tv/cvid/4500191).

40 Cf. Frederick Neuhouser's (2008) study: *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition*.

as they show, for instance, in referring to themselves and their doings as "I ..." or by being aware of their finitude, a fact that entails their concern for themselves, then each of us is likewise able to know that it just won't be appropriate to treat these beings in ways not complying with this understanding of who they are.[41]

Examples of such inappropriate treatment are immediately at hand: using these fellow beings as means serving the satisfaction of my desires without paying attention to their own idea of what is suitable to them; betraying and deceiving them and thus evading their own response to the situation; ignoring their vulnerability in circumstances that do not allow them to help themselves etc. Interestingly enough, we are almost always ready to admit and, what is more, to correct inappropriate performance when dealing with technical matters. When it comes to issues of moral concern, though, we are rather reluctant to admit, let alone correct what we know is inappropriate. Instead, we are prone to offer explanations straight away why we are not able to do what we know we should. Referring to our example once again, it is true, in fact, that each of the prisoners has to bear his burden of grief and despair as a result of a whole system built on corruption, injustice, fraud and crime. It is reasonable, therefore, to say that they will lack the energy to assist others. It is true that "your own misery is always greater, obviously". Even though anybody can perfectly well understand all these justifications for refraining from responding to the suicide's woes, given the circumstances, this kind of indifference, nonetheless, appears to be inappropriate and morally wrong. It goes against what we know ought to have been done even in a situation like this. And so it looks as if Kant is quite right in insisting that we'll have to be aware of our "propensity to rationalize (*vernünfteln*) against

41 In giving importance to human self-knowledge as the source of moral obligation, I am following a main feature of Gerold Prauss's account of practical philosophy. Cf. Prauss 2008. This understanding of moral obligation differs significantly from the so-called "constitutivist" approach, prominently advanced by Christine Korsgaard, which has been much discussed recently (Korsgaard 2009; for a critical analysis of Korsgaard's position, cf. David Sussman's contribution to this volume). Korsgaard argues that "self-constitution", the task of "*making something* of oneself" through rational agency, a task which no human being can escape, is inextricably intertwined with a commitment to morality (Korsgaard 2009, pp. xii-xiii). So it seems we cannot opt out of morality because we cannot opt out of rational agency and engaging in *making something* of ourselves. The approach suggested here carries much less baggage: it does not start out from ideas about an agent's efforts to flourish, to *make something* of him- or herself, to give meaning to his or her life, to meet the requirements of rationality and communication etc. Thus, it avoids the flaw of retreating into arguments that lead to a remarkably thin conception of morality, at best, and/or the flaw of changing the subject by neglecting the objective side of morality which is given with the presence of the Others who may and must demand to be treated according to who they are: beings who are aware of and are an issue for themselves.

those strict laws of [moral] duty and to bring into doubt their validity, or at least their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations".[42]

However, being aware of this weakness and being prepared to resist it does not require one to be a moral saint. Instead, it is about seeing and accepting that the demands of morality, on the one hand, and the demands of our wishes and inclinations, which may include our natural concern for the happiness of our nearest and dearest,[43] on the other hand, do not necessarily coincide. Cases such as those mentioned above show that assumptions to the contrary may rest on wishful thinking. Arguing that we will not be able to gain "true" happiness unless we act in accordance with moral obligations also serves this tendency towards downplaying the fact that meeting these obligations might come at a cost.

Again, it is obvious that we are naturally bound, if not absorbed by our wishes, desires, concerns and the effort it takes to respond to them. Attentiveness and resolution are necessary to take up a stance towards them that allows for taking the – decisive – step back which opens up the scope for recognizing ourselves and others as those who are more than these wishes and desires and their satisfaction and towards whom we have obligations precisely for this reason. The capacity for self-knowledge, entailing self-esteem, self-will and all the other forms of self-relation, turn human beings into very special animals who cannot claim a right to ignorance and innocence any longer. Repression of this knowledge in favour of pursuing our many dealings, bargains, transactions and affairs with cheerful abandon is a permanent temptation. It comes in various guises such as ideologies, rationalizations, and self-deceit. Not least, seriously asking "Why be moral?" may number among them.

III

In the first section of this essay, we have dealt with what has been called the "realities" of human striving, namely human striving for power, and with what might be apt to balance this endeavour. Apparently, some kind of counterpoise is necessary to master its destructive and self-destructive potential and turn it into productive incentives.

42 Kant 1913a, AA 4: 405; my addition. Cf. also Kant 1913b, AA 6: 42.

43 Remember "the unscrupulous judge who believes that the happiness of his own family is justification enough for any such 'Artyoms'". Cf. Khodorkovsky 2014, p. 62 f.

Hobbes sees the point when he compares humans with other living creatures who appear to escape these difficulties and complexities once and for all. If we look at bees and how they live sociably with one another, while having no other directions than their particular assessments and appetites, we might ask, as Hobbes does, "why Man-kind cannot do the same". Of vital importance seems to be the fact that amongst these creatures, the bees, "the Common good differeth not from the Private," whereas man, whose Joy consisteth in comparing himselfe with other men, "can relish nothing but what is eminent".[44] The challenge he faces is to structure his life, which relates to the lives of others in each and every respect, jointly with these others – but in a way that does justice to human self-understanding and self-will. Giving our consent to submitting our wills to the one will of the sovereign – except for a few "liberties" left over because the sovereign won't be able to regulate all our actions –[45], the solution suggested by Hobbes, seems unacceptable. We would consent then to be treated as if we were beasts to be tamed. Hobbes's suggestion demands what a human being cannot grant: to renounce the very qualities by which it is marked human.

Rebuffing the foremost rights of persons, a consequence Hobbes's idea of an almighty sovereign seems to imply, cannot be accepted – in fact for moral reasons. This holds true even though some kind of peace and maybe happiness in return may be expected and even if a contract has been made. Also, it is very unlikely that a Hobbesian-style government could ever "work" in the long term: for, as Alexander Dunlop Lindsay wrote, "there will come a point at which no amount of legal or constitutional machinery will avert disaster".[46] This has to be assumed precisely because this kind of government contradicts morally relevant features of human existence which make themselves felt and cannot be permanently suppressed.

The extreme opposite position was developed by a thinker who shares Hobbes's unsentimental approach and his interest in the phenomenon of power and power relations between humans: Nietzsche. He inveighs against "the social straightjacket"[47] which might have been useful in the early days of mankind when individuals had to get accustomed to fitting in with communities, to abiding by common rules, to being "predictable" so that they could live together. But now, Nietzsche argues, the time has come for the individual to free itself from the "morality of custom" (*Sittlichkeit der Sitte*) and advance towards being an "autonomous supra-ethical individual" (*autonome[s] übersittliche[s] Individuum*). "Au-

44 Hobbes 1991, Chap. XVII, p. 119.
45 Hobbes 1991, Chap. XXI, p. 147.
46 Hobbes 1950, Intro., p. XXX.
47 Nietzsche 1988c, p. 293 (= GM II, 2) / *engl.* Nietzsche 2006, p. 36.

tonomous" and "ethical", Nietzsche adds, "are mutually exclusive".[48] The implication of this statement seems to be that the autonomous individual is not required to accept any (or at least any ethical) restrictions on its striving for whatever goal it pursues for whatever reasons. And this claim, indeed, forms the foundation of Nietzsche's theory of the "will to power".

The individual does not find itself isolated, of course, but has to be seen as part of a whole field of powers in which it has to position itself. When treating power and the will to power, Nietzsche always deals with a *relation* between units that bear a certain strength and force, including intellectual strength. He conceives of power in its different physical, psychological and intellectual facets. As an isolated, absolute power it can never even be thought. It is in need of *an Other* in order to be power at all. Without something to confront it would be empty, as if swinging at nothingness. That is, power needs a resistance, a counter-force against which it reveals and unfolds itself as a power. Consequently, where we speak sensibly of power there is always already posited at least one further power as a counter-force. Accordingly, Nietzsche writes: "The will to power can only manifest itself against resistances; it seeks for that which stands against it."[49]

Even though *the Other*, a counterpart of the individual, is explicitly acknowledged, the framework of Nietzsche's theory allows for all sorts of imbalance. Nietzsche advocates the right of the stronger one to flourish, possibly at the cost of others. He does so because he has his eyes on what he calls the "highest potential power and splendour of man as species" (*die höchste Mächtigkeit und Pracht des Typus Mensch*).[50] They might only unfold in outstanding individuals.

Seeing, in contrast to Hobbes, the promising sides of man's thirst for power, Nietzsche develops a conception of justice nonetheless. It originates, he claims, "among approximately equal powers".[51] "When there is no clearly recognizable supreme power and a battle would lead to fruitless and mutual injury, one begins to think of reaching an understanding", Nietzsche argues. One starts negotiating the claims on both sides, "and so the initial character of justice is barter". Each satisfies the other in that each gets what he values more than the other. "Each man gives the other what he wants, to keep henceforth, and receives in turn that which he wishes. Thus, justice is requital and exchange on the assumption of approximately equal positions of strength."[52] Hence justice, according to

48 Nietzsche 1988c, p. 293 (= GM II, 2) / *engl.* Nietzsche 2006, p. 37.
49 Nietzsche 1988d, p. 424 (= N 9[151]).
50 Nietzsche 1988c, p. 253 (= GM, Vorrede 6) / *engl.* Nietzsche (2006), p. 8.
51 Nietzsche 1988b, p. 89 (= MA I, 92 / HH I, 92) / *engl.* Nietzsche (1986), p. 49.
52 Nietzsche 1988b, p. 89 (= MA I, 92 / HH I, 92) / *engl.* Nietzsche (1986), p. 49.

Nietzsche, is born out of prudence and a sense for the "realities of human striving".

Of course there are and can be, as we all know, quite different relations, relations between humans of very unequal powers and strengths. This fact figures prominently in the background of Artyom's story, which has been discussed above. In cases like this, Nietzsche's idea of justice, which is built on the idea of balancing powers, would not take effect. It simply would not be applicable, since one party might not be able to bring any weight in terms of authority, influence, goods etc. to bear. Following Nietzsche, in these cases there wouldn't be anything available that could "bridle the relentless will to power" (*dem rücksichtslosen Willen zur Macht einen Zaum anlegen*).[53]

However, if human beings are able to relate to themselves by way of knowing themselves, thus being an issue for themselves and demanding to be treated accordingly, the approach Nietzsche suggests would be utterly inappropriate. For then, if this picture is correct, the idea of justice and doing justice to one another could not rely on the equilibrium of power humans possess or are assumed to possess, but would need to be immediately connected to these human features of self-understanding and self-care which have to be respected as such. Hence, doing justice to one another could be only grounded on the idea of strict equality, which Nietzsche rejects, and not on the idea of the equilibrium of power. Nietzsche's alternative contractualist theory on which his account of right and rights relies seems to fail as Hobbes's theory did (but for different reasons). The boundaries of pursuing their aims that individuals have to accept in regard to their dealings with each other (apart from those limits that are set by nature) seem to follow from the internal constitutions of these others, rather than from external conditions such as their actual strength that are both contingent and unstable.

IV

Of course we can say the world we know does trust the "realities" of power relations, on the large and small scale, in politics and in private circles. So perhaps we should put up with these realities and try to get along. We should learn from analyses such as Hobbes's and Nietzsche's. But again, if we are animals that not only want things and pursue them, but know this about ourselves and others, knowing that we are always an issue for ourselves along with all the issues

53 Nietzsche 1988d, p. 221 (= N 5[82]).

with which we are concerned, then we are not innocent and should respect one other as those "ends in themselves" that we are. Accordingly, it won't suffice to be just prudent combatants, skilled in the calculus of power. Furthermore, this understanding should bear on all areas of life. The contrast between the all too often very sad "realities" of human conduct and morality, the divide from which we started out, should be lessened. Why be moral? In order to live up to our knowledge of who we are.

Bibliography

Dreier, James, ed. (2006), *Contemporary Debates in Moral Philosophy*. Malden, MA/Oxford: Wiley-Blackwell.

Gauthier, David (1969): *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford: Oxford University Press.

Gauthier, David (1990): "Thomas Hobbes: Moral Theorist". In: *Moral Dealing: Contract, Ethics, and Reason*. Ithaca and London: Cornell University Press, pp. 11–23.

Hobbes, Thomas (1950): *Leviathan* (1651), ed. by Alexander Dunlop Lindsay. New York/London: E. P. Dutton and Company (= Everyman's Library 691 A).

Hobbes, Thomas (1991): *Leviathan* (1651), ed. by Richard Tuck. Cambridge/New York/Melbourne/Sydney: Cambridge University Press.

Kant, Immanuel (1913a): *Grundlegung zur Metaphysik der Sitten* (= Akademie-Ausgabe, vol. 4). Berlin: Reimer (De Gruyter), pp. 385–464. [*Engl.* Mary Gregor and Jens Timmermann, transl. and ed., (2012): *Groundwork of the Metaphysics of Morals*. Cambridge/New York: Cambridge University Press.]

Kant, Immanuel (1913b): *Die Religion innerhalb der Grenzen der bloßen Vernunft* (= Akademie-Ausgabe, vol. 6). Berlin: Reimer (De Gruyter), pp. 1–202. [*Engl.* Allen Wood and George Di Giovanni, transl. and ed., (1998): *Religion Within the Boundaries of Mere Reason: And Other Writings*. Cambridge/New York: Cambridge University Press, pp. 31–192.]

Khodorkovsky, Mikhail (2014): *My Fellow Prisoners*. London: Penguin.

Korsgaard, Christine M. (2009): *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.

Neuhouser, Frederick (2008): *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition*. Oxford: Oxford University Press.

Nietzsche, Friedrich (1988a): "Unzeitgemässe Betrachtungen: David Strauss der Bekenner und der Schriftsteller". In: Giorgio Colli and Mazzino Montinari (eds.): *Kritische Studienausgabe* (= KSA 1). Berlin/New York: De Gruyter, pp. 157–242.

Nietzsche, Friedrich (1988b): "Menschliches, Allzumenschliches". In: Giorgio Colli and Mazzino Montinari (eds.): *Kritische Studienausgabe* (= KSA 2). Berlin /New York: De Gruyter. [*Engl.* R. J. Hollingdale, transl. and ed., (1986): *Friedrich Nietzsche: Human, All Too Human*. Cambridge, New York: Cambridge University Press.]

Nietzsche, Friedrich (1988c): "Genealogie der Moral". In: Giorgio Colli and Mazzino Montinari (eds.): *Kritische Studienausgabe* (= KSA 5). Berlin/New York: De Gruyter, pp. 245–412.

[*Engl.* Keith Ansell-Pearson, ed., and Carol Diethe, transl., (2006): *On the Genealogy of Morality and Other Writings*. Cambridge, New York: Cambridge University Press.]

Nietzsche, Friedrich (1988d): "Nachgelassene Fragmente 1885 – 1887". In: Giorgio Colli and Mazzino Montinari (Eds.): *Kritische Studienausgabe* (= KSA 12). Berlin/New York: De Gruyter.

Nietzsche, Friedrich (1988e): "Nachgelassene Fragmente 1887 – 1889". In: Giorgio Colli and Mazzino Montinari (Eds.): *Kritische Studienausgabe* (= KSA 13). Berlin/New York: De Gruyter.

Prauss, Gerold (2008): *Moral und Recht im Staat nach Kant und Hegel*. Freiburg/München: Alber.

Rawls, John (1999): *A Theory of Justice*. Revised Edition. Cambridge/MA: Harvard University Press.

Sorell, Tom (2007): "Hobbes's Moral Philosophy". In: Patricia Springborg (Ed.): *The Cambridge Companion to Hobbes's Leviathan*. Cambridge/New York: Cambridge University Press, pp. 128 – 153.

Tralau, Johan (2007): "Leviathan, the Beast of Myth: Medusa, Dionysos, and the Riddle of Hobbes's Sovereign Monster". In: Patricia Springborg (Ed.): *The Cambridge Companion to Hobbes's Leviathan*. Cambridge/New York: Cambridge University Press, pp. 63 – 81.

Weber, Max (1980): *Wirtschaft und Gesellschaft: Grundriss der verstehenden Soziologie* (1922), ed. by Johannes Winckelmann. Tübingen: Mohr-Siebeck. [*Engl.* Günther Roth and Claus Wittich. Eds., (1978): *Economy and Society: An Outline of Interpretive Sociology*. Berkeley/Los Angeles/London: University of California Press.

David Sussman

Morality, Self-Constitution, and the Limits of Integrity

> One must indeed grant something even more unpalatable; that, from the highest biological standpoint, legal conditions can never be other than exceptional conditions [...]. A legal order thought of as sovereign and universal [...] would be a principle hostile to life, an agent of the dissolution and destruction of man, an attempt to assassinate the future of man, a sign of weariness, a secret path to nothingness. —
> Nietzsche, *On the Genealogy of Morals* II, §11

Many philosophers have hoped to show that commitment to moral principles is a necessary condition of being a rational agent. If successful, such "constitutivist" approaches would manage to avoid a dilemma that is often thought to beset any attempt to provide a foundational justification of morality. In "Does Moral Philosophy Rest on a Mistake?", H.A. Prichard (1968) famously argues that there is something profoundly wrongheaded about such efforts. Supposedly, these attempts must either reduce morality to some non-moral concerns (such as those of self-interest), or appeal to considerations that already presuppose some sort of ethical commitment. Prichard contends that because any attempt to provide a basic justification of morality must in this way either change the subject or beg the question, the endeavor is itself ill-conceived. He concludes that we can only reason about moral demands from within the moral point of view, since there could be no external perspective from which those demands might be justified without completely effacing their character as moral demands.

We might avoid Prichard's dilemma by refuting a presumption of both its horns. Prichard's dilemma assumes that there are coherent practical outlooks external to morality that an agent might adopt instead of the moral point of view. What the constitutivist hopes to show is that there really is no such thing as a wholly non-moral perspective that an agent could hope to inhabit, and so that the moral skeptic really has no place from which he can issue his challenges. The constitutivist does not have to deny that there are non-moral interests that are distinct from those of morality. She need only claim that there is no way of being an agent who self-consciously pursues any kind of interest who does not implicitly recognize the supreme authority of moral principles. The constitutivist agrees with Prichard that the only perspective from which we can consider moral claims is from "within" the moral point of view, but denies that there is any space outside that sphere that could serve as its boundary. The demand for a non-moral justification of morality is thus not ignored, but neither is it satisfied on its own terms. Instead, the constitutivist aims to show that this demand

rests on the illusion that there is some real alternative to being a morally committed agent that is open to us, an alternative we need some special reason to reject.

To reach this conclusion, some philosophers argue that an agent's psyche or "soul" must have a special kind of unity or integrity that can only be maintained through the embrace of morality. In the *Republic*, Plato contends that for people or polities to be capable of acting, their various constituents must recognize a certain structure of laws that determine the proper relations of authority and deference between them. Such laws serve to define the difference between when the agent is herself truly acting, and when she is merely being affected by some forces from within or without herself. Aristotle similarly holds that the form of life characteristic of human beings is one governed by reason, which determines what is to count as the healthy functioning of basic human powers in relation to each other. For both Plato and Aristotle, the more vicious or unjust a person is, the less she really is a person, and the less she can really do anything. The bad person turns out to be capable of action in only a partial or derivative sense, in something like the way that being sick or dying still counts as a way of living.

Kant also argues that recognition of a basic principle of reason is needed for a being to have any power to act in the world in a way distinct from the operations of natural causes. Unlike Plato and Aristotle, Kant does not hold that a fundamental function of practical reason is to apprehend substantive truths about goodness and value to guide the proper functioning of the soul. Kant instead argues that the fundamental moral law takes its content merely from what would count as success in properly establishing a will in the first place. The practical reasoning that defines the will must then be autonomous; such thought would have to proceed simply from a commitment to free self-determination, simply as such.

Unfortunately, all of these approaches bring a great deal of metaphysical baggage with them. Plato and Aristotle both presuppose an order of substantive ethical truths about goodness and excellence that are epistemically accessible to us in ways we can trust. Kant avoids these problems, but relies on a notion of free self-determination that is hard to reconcile with our also being parts of the natural world. In this paper, I focus on a modern version of constitutivism that, while drawing heavily on Plato, Aristotle, and Kant, proceeds in a way intended to be fully consistent with contemporary naturalistic scruples. In recent work, Christine Korsgaard (2008; 2009) has argued that the basic function of action as such is to produce and sustain just that unity that makes agents capable of acting. Supposedly, the formal aim of action is to integrate agents in a way that enables them to act, much as the proper functioning of any living thing serves to sustain and reproduce that organism in just those ways that allow it

to continue in this very activity of living. For Korsgaard, the self-constituting function of action shows that practical reasons must be "public" or "shareable" in a way that grounds not only principles of instrumental and prudential rationality but also a recognizably Kantian universalizability requirement.

I consider Korsgaard's attempts to derive morality from integrity in light of challenges from Bernard Williams, who largely echoes Nietzschean criticisms of Kant. For Williams, integrity is indeed central to agency or "character," but such integrity is ultimately in tension with the kind of impersonal universality that Korsgaard hopes to extract from it. Following Williams, I argue that the ideal Korsgaard develops is a description not of fully rational agency, but of a kind of compulsive self-denial, something akin to Nietzsche's "bad conscience." Paradoxically, Korsgaard's fully realized agent is one that identifies with nothing other than his own alienation from himself, and it is only such pervasive alienation that gives his will the appearance of an especially durable sort of unity. However, this result does not require that we abandon all hope of a constitutivist defense of morality. In conclusion, I offer a different way of understanding the essential integrity of agency that, although more modest and less alienated than Korsgaard's, may still prove adequate to ground the main principles of Kantian ethics.

1. Korsgaard holds that the problem of integrity is engendered by the basic powers of rational reflection that we exercise whenever we deliberate about what to do. This problem is not faced by animals who unreflectively inhabit their natural teleology, and cannot wonder whether things really are as they seem, or whether they should satisfy some desire they happen to feel. The behavior of non-rational animals seems to be immediately determined by their instincts and perceptions, which under normal conditions produce activity that promotes the continued good functioning of that organism. Although an animal does not have any substantial self-conception, it does see the world egocentrically, through the lens of its needs. Such a world appears not as a set of publicly available objects with general causal powers, but as an array of resources, opportunities, and threats for that animal (2009, pp. 109 – 111).

In beings like us, however, rational reflection makes it possible for what we perceive to diverge from what we believe, and for what we desire most to come apart from what we actually intend to do. From such a reflective perspective, my desires and perceptions cease to be immediate causes of my thought and behavior, becoming instead something more like proposals for how I am to think or act. At this point, I need to form some conception of the objects of my desires that will allow me to distinguish between more or less successful ways of pursuing them. In confronting the manifold of my wants, I face the further question of

which desires I really care about, and which are to be treated as merely distractions or temptations to be resisted, despite their motivational influence. In this way, reflection forces an agent to confront an issue that is not faced by an animal: the question as to how far some feature of her mental life counts as her own doing, and to what extent that feature is merely a condition from which she suffers and must somehow work around.

Reflective agency might seem to require only that we add to the animal an ability to pick some desire to pursue from all those it experiences. In "Freedom of the Will and the Concept of the Person" Harry Frankfurt (1988, pp. 11–25) argues that what is essential to being a reflective agent is the ability to form and act from "second-order volitions": that is, higher-order desires that some other desire be the one acted upon. Yet it is hard to see how just adding another desire with a peculiar content serves to deal with the alienation and fragmentation of the self that reflection creates. These problems concern how to distinguish responses in which the agent counts as truly doing something from those in which she is passively affected by something within her. If this question can apply to ordinary first-order desires, it would seem equally applicable to higher-order desires as well. Just as we might have any number of first-order desires, we might equally have many different second-order volitions. It will hardly help to introduce still higher-order desires that a particular second-order volition be the one that determines our behavior, since the same problems will just reinscribe themselves at each new level.[1]

Frankfurt responds to these worries by arguing that our agency involves not just the formation of higher-order desires, but the exercise of a special power of identification whereby the agent commits herself to act on a desire in a way that "resounds through" and harmonizes the rest of her motivational economy. Frankfurt contends that identification is not just the creation of another desire with some special content, and for him it is not logically possible for a person to be alienated from her own identifications or passive before them. For Frankfurt, such identification comes to be the defining power of rational agency; the agent is just that which so identifies, and its actions are just those behaviors that express such identifications.

The problem with this response is not so much that it is false but that it is unilluminating. On Frankfurt's account, identification functions much like a dormitive virtue. Frankfurt manages to distinguish a person being truly active with respect to her psychic life simply by postulating a *sui generis* power of real self-determination, which is defined so that the question of alienation cannot even

1 See Watson (1990).

arise about it. This move does not solve the problems of unity and activity but only relocates them: *what* could it be to identify with some desire or attitude such that this identification would be so intimately mine that I could never experience any kind of estrangement from it?

A central insight of constitutivism is that for any kind of thinking to count as the active exercise of a power, such thinking must be situated in a broader context of normative principles or laws that give it its significance. Moving pieces of wood on a board can count as castling only if a game of chess is being played; that is, only if this movement is performed by someone who takes her actions to be governed by the relevant sorts of rules, in the context of a broader functioning practice that others participate in. Similarly, entering into a contract or casting a vote is something more than just uttering special phrases or making marks on paper with certain intentions or beliefs. For any such doings to be contracting or voting, the requisite legal and political institutions must be in place, and those engaged in these activities must recognize those practices as in some way authoritative over what they are doing.

Constitutivism holds that for any response to a desire to count as a decision to act, that response must be situated within some formal principles of choice, which define both what counts as truly making a decision and what properly follows from doing so. If nothing else, a reflective agent must take her choice to be governed by some sort of instrumental norms. In adopting any particular end, the agent has to take for granted that she has a general power to bind or obligate herself to do whatever is requisite to bring about that end, even if such action is not immediately or independently appealing to her. The imperative to take the necessary means to our ends can be understood as a description of a basic normative power any agent must ascribe to herself in order to see herself as an agent. In deciding to do something, an agent commits herself to being a certain kind of cause in the world, and so sees herself as having the power to bind herself to doing whatever is required for such efficacy.

The authority of instrumental norms serves to give a manifold of psychological states and physical movements the distinctive unity of an action, allowing for a normative interpretation of that behavior in terms of varying degrees of success and failure, perseverance or weakness. Yet we need more than merely instrumental norms to make sense not just of actions, but of the distinct, enduring agents who might be performing them. Such an agent would have to have an identity that can go beyond any of its deeds. This identity is needed to ground the possibility that the very same agent might not have performed that particular act, or that she might go on to perform different acts as well. Instrumental principles establish how different behaviors might all be parts of the same pursuit of an end, but such principles do not show how different choices of ends could all

be the decisions of one and the same agent. Yet without such non-instrumental principles, we cannot make sense of how an agent might change her mind (or decline to do so), since both change and constancy presuppose the possibility that the very same agent could still persist even if she came to abandon the end she has adopted.

In *Self-Constitution*, Korsgaard argues that in addition to the hypothetical imperatives that individuate acts, an agent must recognize "categorical principles" that determine to what extent different choices of ends still count as his own doing. She explains:

> The hypothetical imperative binds you because *what* you are determining yourself *to be* when you act is the *cause* of some end. The categorical imperative binds you because *what* you are determining to be *the cause of* some end is *yourself*. (2009, p. 81, Korsgaard's emphasis)

At this point in her argument, the "categorical imperative" does not have any distinctive content. The term is really just a placeholder for whatever principles govern not the relation of means to ends, but how the agent goes about adopting ends in response to a manifold of incentives in the first place. Such principles make up the "constitution" of an agent's soul. On this view, an agent is not to be identified with any feature of her constitution to the detriment of others; neither reason nor emotion nor any other part is taken to be the agent's "true self." Instead, the agent supposedly identifies with and is defined by the entirety of her constitution, and it is this whole structure that determines how far some behavior is something she does, and to what extent it is something that merely happens to her and for which she may be less than fully responsible.

2. Following Plato, Korsgaard considers five different political models for how an agent might constitute herself: aristocracy, timocracy, oligarchy, democracy, and tyranny. The aristocratic constitution is supposedly highest, being governed by reason in a way that realizes a form of Kantian autonomy. According to this constitution, the agent's appetites and inclinations serve to propose various objects and courses of actions to the soul as a whole, which are to be judged according to universal rational principles that are prior to all such desires. Korsgaard departs from Plato in holding that this ultimate standard is to be found not in some independent order of normative facts, but instead in the basic task of sustaining the very unity of the agent that makes action possible. The aristocratic soul only recognizes a desire as providing a reason for action insofar as it can be understood as a fully public reason; that is, in something like fair terms

that can be similarly appreciated by that agent or others in different circumstances, when a very different array of desires or moods might come into play.

For Korsgaard, the more specific conceptions of the aristocratic soul correspond to different interpretations of the scope of this publicity requirement. If a good reason only needs to be sharable by the various perspectives within a natural individual's biography, then the aristocratic soul is a rational egoist who, although indifferent to other people, regards all different times in his life as being equally significant. If the standard of publicity is expanded to include all agents simply as such, then the aristocratic soul becomes the Kantian agent who fundamentally conceives of herself as a "citizen of a universal kingdom of ends." Such a person is dedicated to acting only on principles that could be freely accepted by all others who take themselves to have an equal say in and responsibility for the governance of their common life.[2]

Closest to the aristocratic constitution is the timocratic soul, where the agent is governed by a concern for his integrity in the more familiar sense of being honorable, respectable, or "up-standing." As Korsgaard presents it, the timocrat is committed to acting in ways that are recognizably "fine" or "noble." He is dedicated to preserving his honor in a way that might be recognized by others who are themselves taken to be worthy of such esteem. The timocratic agent is governed by reason to some extent, but a reason that is not fully distinct from pride or self-love. The honor-lover is unable to conceive of a standard of correctness that is independent of the evaluative attitudes of a particular kind of ethical community. Such an agent cannot countenance the possibility that the right thing to do might be something that no one he knows or respects could appreciate.

Unlike the timocrat, the oligarchic soul does not bring any ideal of individual excellence to bear on her desires. Instead, the oligarch has certain dominant passions that she takes as her real needs simply because of their motivational strength, and evaluates the acceptability of other desires in terms of their compatibility with these powerful concerns. She thus operates from a "present-aim" conception of prudence, unlike the temporally neutral self-concern that characterizes the aristocratic rational egoist. In the absence of a stable set of dominant

2 In *Self-Constitution* (2009), Korsgaard does not focus on the question of why the aristocratic soul should take the Kantian form rather than merely that of the rational egoist. Korsgaard sees rational egoism as one proposal for how an aristocratic soul might constitute itself, a proposal that cannot be assessed in the absence of a defense of some more specific understanding of what that agent's overall self-interest involves. Korsgaard's doubts that such a defense is possible without presupposing the broader commitments of Kantian autonomy are developed in "The Myth of Egoism", in Korsgaard (2008), pp. 69 – 100.

passions, an oligarchic soul may decay into a democratic one, in Plato's peculiar sense of "democracy." The democratic soul resolves to pursue whatever desire happens to be most salient at the time, as if by a kind of motivational lottery. In essence, his fundamental principle is to act as if he had no principles at all.

Like Plato, Korsgaard considers the most degenerate soul to be that of the tyrant. Tyranny is what emerges from a democratic constitution when one passion attains such power that the question of whether to act on it cannot even be seriously considered. The tyrant is unable to really deliberate about her basic ends at all; she cannot make up her mind about what to pursue, because with regard to basic ends her mind is always already made up. As Korsgaard describes it, the tyrannical soul is like that of someone in the grip of an addiction or sexual obsession, who acts not from commitment but from compulsion. Like the aristocrat, the tyrant manifests a strong kind of practical unity. Unfortunately, this unity is really just that of single desire that has come to be the tyrant's entire self, precluding any real reflection or deliberation about other ends. Since she cannot choose her ends, the tyrant cannot make sense of herself as a distinct agent with the power to bind herself to even instrumental requirements of action. Korsgaard concludes that the tyrannical person's doings are not interestingly different from animal behavior that is immediately determined by instinct. On her view, the tyrant is not so much a defective agent as the last remains of one, being all that is left of a will after a process of corruption and dissolution has fully run its course.

3. Korsgaard contends that the forms of agency that lie between aristocracy and tyranny are all functional but inherently defective. Although these people can act, they do so in only a derivative or degenerate way. Korsgaard holds that the common flaw of these souls is that their integrity is objectionably contingent. These agents can expect to maintain their unity only so long as certain kinds of practical problems never happen to arise. In contrast, the rationally-governed soul of the aristocrat can recognize itself in any conceivable situation that she might confront:

> On certain occasions, the people with the other constitutions fall apart. For the truly just person, the aristocratic soul, there are no such occasions. Anything could happen to her, anything at all, and she will still follow her own principles—and that is because she has universal principles that can consistently be followed in any kind of case...She is completely self-possessed; not necessarily happy on the rack, but *herself* on the rack, herself even there. (2009, p. 180)

Supposedly, the aristocratic or autonomous agent has a special kind of integrity that shields her from luck in a way that makes her actions more truly her own

than those of the other, less reflectively structured kinds of soul. Korsgaard concludes that the other forms of agency are not really self-conceptions that an agent could endorse instead of an aristocratic one. These other constitutions only describe what a person will have ended up being should her efforts to fully integrate herself fail in various ways. Supposedly, no one can coherently choose to be such a non-autonomous agent, even if that is all that they might turn out to be in the end.[3]

Korsgaard is not making the mistake of thinking that because our souls must possess some degree of durability, any decrease in fragility must be an improvement.[4] Neither does she deny that even aristocratic souls might break down when subjected to enough pain, fear, or disease. Her point is that in such cases this disintegration would be merely something that *happens* to the soul, effected by external forces that overwhelm and shatter it. This threat is different from the possibility that the agent might come to actively dissolve herself, in a way that counts as a proper expression of her fundamental commitments. For Korsgaard, what is distinctive about the aristocratic soul is that it alone is immune to turning against itself in this way. All souls may break down, but the non-aristocratic ones also stand ready to actively undo themselves should the right conditions arise. Korsgaard seems to see this not merely as failure, but as something more like self-betrayal. After all, once I am prepared to betray you under certain conditions, it would seem that I have already done so, even if those conditions never arise.

For contrast, consider the timocratic person. This agent is governed by pride or love of honor, wanting above all to be recognizably good, fine, noble, etc. The timocrat aims at both merit and the appearance of merit, and so faces a dilemma in contexts where these desiderata diverge. The unlucky timocrat could find himself in a situation where the only right acts are profoundly distasteful ones (e. g., retreat in battle, surrendering a friend to the authorities), so that it would be im-

3 Here Korsgaard can answer David Enoch's charge that even if some moral principle can be shown to be constitutive of a "game" that all agents must be playing, these agents might still be merely "going through the motions grudgingly, refusing to internalize the aims of the game." (Enoch 2006, p. 189.) Even if an agent ends up merely "going through the motions," this is not what she can set out to do, or what she fundamentally takes herself to be doing as she so acts. An agent can be intentionally going through the motions of some game only insofar as she sees this as a way of doing something else whole-heartedly (such as play-acting, teaching, dissembling, etc.).

4 In his review of *Self-Constitution*, Sergio Tenenbaum worries that Korsgaard has "illegitimately turned unification into a goal to be maximized," observing that "a crystal chalice is more likely to break and lose its unity than a glass one, but it's not a defective chalice for that reason." (Tenenbaum 2011, p. 454).

possible for him to maintain his elevated sense of his own nobility. In the face of such circumstances, the timocratic soul will be unable to identify what he really wants; his basic commitments will impel him to both embrace and reject the same things.

Although the oligarch does not share the timocrat's desire to preserve a privileged image of herself, she too is vulnerable to self-dissolution. The oligarch has certain powerful passions that she takes to define her needs, and evaluates the acceptability of any other desires in terms of their compatibility with these dominant interests. As a result, the oligarch can be confident that she will consistently pursue her projects only so long as the economy of her desires remains stable. Past a certain degree of motivational influence, what had been a temptation will begin to assert itself as a new dominant passion, and those desires that used to dominate will start to appear as the real temptations. During such a transition, the oligarch will be unable to consistently identify what she really wants or chooses; that is, she will be unable to clearly recognize herself in any activity that would count as her own choosing. The situation is even worse for the democratic soul. Unless the overall economy of his desires is unusually stable and coherent (or his environment extremely limited), the democrat can expect always to be flitting from goal to goal as they catch his fancy, and so be unable to engage in any complex or long-term projects (Korsgaard 2009, pp. 168 – 169).

4. The aristocratic or autonomous soul is supposedly the only one that does not face the prospect of undoing itself because the aristocrat is dedicated above all to the activity of sustaining and reproducing her formal unity as an agent. Such a person is committed to acting only on principles that she could still recognize despite any changes in her moods, desires, or social attachments. Supposedly, the one interest every agent must have is in maintaining the integrity necessary for any kind of action at all. By identifying with this supposedly necessary interest, a person manages to constitute herself in a way that will remain available to her to endorse whatever changes in her other concerns and attitudes she might undergo. Korsgaard claims that

> every rational agent must will in accordance with a universal law, because it is the task of every rational agent to constitute his agency. And the law ranges over all rational beings, that is, it commands you to act in a way that *any* rational being could act, *because you could find yourself in anybody's shoes, anybody's at all, and the law has to be one that would enable you to maintain your integrity, come what may.* (2009, p. 214, my emphasis)

The aristocratic soul, by identifying with a standpoint in abstraction from any of its substantive incentives, is the only one able to stand above the fray of contin-

gency that could bring the other souls to willingly dismantle themselves under certain conditions.

Korsgaard contends that only an aristocratic constitution can be coherently affirmed or endorsed by an agent. For her, the other constitutions only describe what results when a person, striving to be autonomous, falls short in some way. Korsgaard explains:

> Plato's argument shows that this aristocratic constitution is the only one you can choose.... *Timocratic, oligarchic, and democratic souls disintegrate under certain conditions, so deciding to be one would be like making a conditional commitment to your own unity, to your own personhood.* And that's not possible. For consider what happens when the conditions that cause disintegration in these constitutions actually occur. If you don't fall apart, have you failed to keep your commitment, like the conditionally just person who holds out on the rack after all? *But if you do fall apart, who is it that has kept the commitment? If you do fall apart, there is no person left.* You can be a timocratic, oligarchic, or democratic person, in the same way that you can be a just person who fails on the rack. But you cannot decide in advance that this is what you will be. (2008, p. 120, my emphasis)

Supposedly, the timocratic person is really committed to being honorable only so long as his merit and the prospects of social recognition do not diverge too far; otherwise, he has implicitly resolved to abandon those ideals of honor, and so cease being himself. Korsgaard claims that it would be logically impossible to live up to such a commitment, because in so doing there would no longer be an agent to whom we can attribute such success ("if you do fall apart, *who is it* that has kept the commitment? If you do fall apart, there is no person left").[5]

Korsgaard's argument turns on the assumption that if we adopt a commitment in the knowledge that it can be pursued only under certain conditions, then we have really only adopted a conditional commitment; that is, we have decided to hold to this commitment only so long as the necessary conditions obtain, implicitly resolving to abandon it when they don't. By way of illustration, Korsgaard considers a person who claims to be devoted to justice, but only so long as he is not put to the rack (in which case he is ready to do whatever is necessary to be released, however shameful or unjust). Korsgaard does seem right about this case, holding that such a person cannot really be dedicated to justice, but must instead have a deeper commitment to avoiding suffering, a concern that sets the conditions in which his subordinate interest in justice operates.

---

5 This last step in the argument would seem to prove too much, insofar as a similar move would show that a person cannot coherently intend to kill himself. After all, if the intention is successfully executed, there will no longer be an agent to whom it could be attributed. Similarly, a parliament could not decide to dissolve, or a band to break up, or a meeting to adjourn.

Korsgaard observes that although a truly just person might yield under torture (and even know that he will), such a person cannot decide in advance to break in this way. To maintain his integrity, the just person does not have to ignore or deceive himself about human frailty. Nevertheless, he must refuse to incorporate those conditions into his own basic practical principles, treating them instead as external forces that may overwhelm him rather than as considerations that might give him reason to change his mind.

5. We might understand integrity in a different way. Korsgaard contends that true integrity involves being able to recognize yourself in any conceivable choice situation, "come what may." For Bernard Williams, this is exactly backward. Williams argues that having character or integrity as an agent depends crucially on the *inability* to recognize oneself in certain kinds of choices. Williams (1981) asks us to consider a man who can save his beloved wife from drowning, or some strangers, but not both. Williams concludes that if this man's readiness to save his wife depended on his awareness that it is morally permissible to do so, he would be guilty of entertaining "one thought too many." For Williams, to require that this man respond in this way would be to demand that he be fundamentally alienated from his deepest concerns in a way inconsistent with the integrity that we need in order to be agents.

Williams's point here is not about the content of morality; he readily accepts that on most plausible moral views, it does turn out to be permissible for this man to save his spouse. Nor is Williams denying that substantive moral demands are indeed authoritative principles of practical reason. At least for the sake of argument, Williams is willing to grant that to act is to exercise the power of practical reason, and that having such a power involves commitment to moral principles, principles that in some cases might require a person to sacrifice his heart's desire. The example is instead meant to address the more basic issue of why we see ourselves as agents in the first place.

Williams contends that we become practically engaged with the world by way of certain "ground projects" or "categorical desires" that give us a reason to keep living, which might often involve deep attachments to particular people or traditions. These projects don't give us a *justification* to live; if any such justifications were needed, they could be readily supplied by any of the worthy projects with which moral reason presents us. Rather, such attachments are meant to give us an *interest* in living, without which we would become dissociated from the world, not caring one way or another what happened to us or anything else. Williams argues that no impersonal rational principles can demand that we abandon the basic concerns that bring us to see ourselves as particular agents in the first place, even if coherent agency requires us to accept those prin-

ciples. For Williams, the very real demands of impersonal practical reason depend upon a profoundly personal, non-rational form of interest in one's own life.

If Williams is right, there is something very wrong about requiring any agent to be able to act only on principles that she could address to herself in any conceivable circumstances, including those in which her ground projects would no longer make sense. Consider the case of Ajax that Williams discusses at some length in *Shame and Necessity* (1993, pp. 72–102). As depicted by Sophocles, Ajax certainly seems to have a timocratic constitution, valuing his honor above all else. Enraged over a perceived slight, Ajax sets out to massacre his compatriots but, deluded by Athena, slaughters only some sheep and their inoffensive shepherds. In so doing, he makes a laughingstock of himself just when he thought he was at his most terrifying. Knowing that he can never be taken seriously as a warrior again, Ajax takes his own life.

Should we conclude that Ajax had implicitly resolved to live an honorable life only so long as he could gain public recognition, and to kill himself if he could not? If so, then Korsgaard would be right to see Ajax as never really being dedicated to honor in the first place, being more concerned with the esteem and the admiration of his peers. The apparent unity of Ajax's character would then be an illusion sustained by his systematically deceiving himself about his own commitments, which is certainly a real human possibility. However, Williams suggests an alternative that Korsgaard cannot so readily accommodate:

> [Ajax] knows that after what he had done, this grotesque humiliation, he cannot live the life his *ethos* demands.... Being what he is, he could not live as the man who had done these things; it would be merely impossible, in virtue of the relations between what he expects of the world and what the world expects of a man who expects that of it. (1993, pp. 72–73)

If Williams is right, Ajax could neither have explicitly nor implicitly resolved to abandon his commitment to leading an honorable life should he be thoroughly humiliated, precisely because such humiliation makes it impossible to continue to lead that life, regardless of anything he might decide.

It may well be that after such a disgrace no way of going forward could count as an expression of a commitment to Ajax's particular code of honor, and so no response would be a way of holding fast to that code. Ajax would then be unable to confront a question of whether to abandon or retain his practical identity because that identity would simply no longer be available to him after such a catastrophe. His suicide could not express a choice to no longer go on in such circumstances; instead, it would only be an acknowledgement that there was no longer any way for *him* to go on, given what happened. A life without honor

is not a life Ajax could care about or recognize as his own, and even if he thought he might come to take an interest in such a life, he nevertheless has no interest in becoming such a person.

Similar examples might be found closer to home.[6] I might suffer a fate like Ajax's if forced to choose which of my children to save from death, or whether to care for my ailing mother or fight fascism. In all these cases, there may be no way I could make such a choice that would allow me to still recognize someone as myself at the other end. Of course, an agent might come to find such a way in any particular case, or at least arrive at a new understanding of herself that would allow her to live even after these crises. But if I found myself unable to do so, would this show that there is something defective in my basic commitments, something that keeps me from being a fully realized agent?[7] Would I be making progress if, in some cool hour, I could make up my mind about which of my children to save should I ever be forced to choose between them?

These dilemmas differ significantly from Korsgaard's just man facing the prospect of torture. Although that man knows that it is highly unlikely that he would be able to remain himself throughout such an ordeal, he can still form some conception of what it would be for him to do so. The just man might foresee that he will completely break down under torture, but this prediction itself presupposes an understanding of what it would have been for him to remain whole that is needed for the idea of his disintegration or dissolution to make sense. Torture threatens only the just man's ability to live up to his principles, rather than their meaning or applicability to himself and his world. In contrast, in contemplating the choice of which of my children to save I can form no real conception of what could count as remaining myself through such a dilemma, which would require a profound betrayal of someone I love unconditionally. I cannot begin to imagine how to relate to myself in such a future, or how such a future person could make sense of his past in a way that includes me as I am now.

Korsgaard may well be right that to be unified enough to act, I must embrace principles that determine how I should relate to myself over the entire range of

---

6 Following Korsgaard, I here only consider dilemmas that might afflict individuals. However, the most compelling examples will probably involve the collapse of a whole way of life and the ethical concepts that it served to define. For an illuminating discussion, see Lear (2006).
7 Cf. Williams: "People do not *have* to think that they could not live in that situation; they do not *have* to think any such thing, and this is a type of ethical thought as far removed as may be from the concerns of obligation. But they may sensibly think it if their understanding of their lives and the significance their lives possessed for other people is such that what they did destroyed the only reason they had for going on." (1993, p. 74, Williams's emphasis).

real possibilities for who I might become. However, there seems to be no need for such principles to include options that would go beyond what I could imagine being without losing my grip on myself and those with whom I might interact, even if I admit that such possibilities might be successfully realized by someone else.[8] The idea of a principle that would allow me to recognize myself under any conceivable circumstances makes no more sense than that of instincts that would enable an animal to flourish under any kind of environment, no matter how bizarre or unlike its natural habitat. The only way of being assured of remaining oneself "come what may" is by refusing to become anyone in particular, just as the only way to escape the threat of death is by never coming to life in the first place.

6. Fortunately, constitutivism may not require as strong a sense of the unity of agency as Korsgaard supposes. I have argued that the insistence that we be able to maintain our integrity "come what may" does not merely require that we be partially alienated from any particular interest we may find ourselves with. In addition, this requirement would seem to make such alienation the constitutive principle of agency; the true agent would be fundamentally defined by her refusal to fully identify herself with any substantial concern or attachment. In so doing, Korsgaard leaves us unable to see less reflective and more wholehearted forms of life as anything other than defective forms of the ironic consciousness characteristic of modern life.

However, Korsgaard is right in thinking that any effort to set the limits of one's self by choice must be self-defeating. For example, I might decide to make Judaism central to my identity, refusing to recognize any future person who lacked these religious commitments as a potential version of myself. Such an effort would have to miscarry because, in order for this choice to even be available to me in the first place, I would have to be able to envision a real possibility of a non-Jewish future of myself, if only to be in a position to repudiate it. As Korsgaard suggests, attempting to legislate the limits of who one might be is like trying to hear claims made in a language one understands as mere noise, or refusing to see the emotion in someone's facial expressions.[9] Such efforts can only appear to succeed through a massive exercise of self-deception, where the

8 Here I draw on the distinction between "real" and merely "notional" confrontations that Williams (1985) introduces in the ninth chapter of *Ethics and the Limits of Philosophy*. See also Frankfurt (1988), pp. 177 – 190.

9 "I am saying that responding to another's reasons as normative is the default position—just like hearing another's words as meaningful is the default position. It takes work to ignore someone else's reasons..." (2009: 202). See also (1996: 139 – 144).

agent convinces herself that she cannot grasp that which threatens her precisely because of what she sees it to be all too clearly.

I have argued that the integrity needed for action requires only that the agent embrace principles that can be shared with not every human possibility, but only those possibilities in which she could see herself and the social world around her as something she could actually inhabit and in which she could find her way about. Yet although there is a significant metaphysical distinction between these conceptions of agency, that distinction may not have much bearing on first-person deliberation. The difference between what I could be and what somebody like me could be depends on the bounds of my practical imagination, and of my ability to recognize myself in some possibility in a way that still engages distinctive forms of self-concern. Although there must be some such limits, no rational agent can ever be very confident about just where they lie in her own case. The problem is not just that we may be confused, biased, or self-deceived. In addition to these familiar failings, the very attempt to discern the bounds of my practical possibilities may change what those limits are, turning what had been merely notional possibilities into real options that can tempt or threaten me in ways that were previously inconceivable.

A properly integrated agent may still need to extend her principles as far as they can go, and to live by maxims that are as universalizable as possible. This need does not presuppose that the agent thinks that she could really turn out to be anybody, but only that there can be no determinate bounds to her possibilities prior to her best efforts to discern and realize them. Kant's moral law could then be recovered as a regulative rather than a constitutive principle of rational agency. So understood, that law would be grounded in our basic task of establishing the limits of ourselves by seeing just how far we can share or publicize our practical reasons and still recognize ourselves in a social world that we could hope to navigate. There must be some limits to any such efforts, but we do not have to try to identify or impose them in advance; the world itself will do that without any help from us. Our striving to constitute ourselves may then proceed just as if there were no such limits, since there is no danger that we might actually overstep them. In doing so I would be discovering not what rational agents as such could will, but rather the extent to which some possible ways of living are real options for me in particular. If so, then Kant's seemingly impersonal ideal of citizenship in a universal kingdom of ends could be recovered as an honest affirmation of the contingency and fragility of our particular identities.

7. I have argued that it is a mistake for the constitutivist to try to derive substantive moral principles from a perfectly generic idea of action or rational agency. Instead of directly replying to the challenge "Why be moral?", the constitutivist

should address what is invariably the real worry: Why should *I* be moral? In response, we can show that by merely advancing this question in the first-person, the challenger reveals a commitment to some principles of obligation and interpersonal accountability. The scope of these principles will correspond to the forms of social life that the agent could inhabit, and so to the range of possibilities of who she could become or otherwise relate to as a person. The idea of a rational agent simply as such will then come into play not because there can be no limit to the self's possibilities, but because those limits only come to be what they are through a subject's attempts to exceed them. If successful, the constitutivist's reply would show that the question "Why should I be moral?" is really just a way of asking why I should be at all, and so recognize any actions or decisions as interestingly *mine* in the first place. Perhaps it is possible for a person to become so detached or alienated as to find himself really in need of an answer to that query. We may indeed have nothing helpful to say to such a profoundly dissociated being. However, this lack should not undermine the moral confidence of the person who cannot help but approach the world from a groundless concern for some life as essentially and distinctively her own.

Bibliography

Enoch, David (2006): "Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Action". In: *Philosophical Review* 115, no. 2, pp. 169–198.

Frankfurt, Harry (1988): *The Importance of What We Care About.* Cambridge: Cambridge University Press.

Korsgaard, Christine M. (2009): *Self-Constitution: Agency, Identity and Integrity.* Oxford: Oxford University Press.

Korsgaard, Christine M. *The Constitution of Agency* (2008): *Essays on Practical Reason and Moral Psychology.* Oxford: Oxford University Press.

Korsgaard, Christine M. *The Sources of Normativity* (1996): Cambridge: Cambridge University Press.

Lear, Jonathan (2006): *Radical Hope: Ethics in the Face of Cultural Devastation.* Cambridge, MA: Harvard University Press.

Nietzsche, Friedrich (1989): *On the Genealogy of Morals*, Walter Kaufmann, trans. New York: Vintage Books.

Prichard, H. A. (1968): "Does Moral Philosophy Rest on a Mistake?" In: *Duty and Interest.* Oxford: Oxford University Press, pp. 1–17.

Tenenbaum, Sergio (2011): "Review of Christine Korsgaard's *Self-Constitution*". In: *Ethics* 121, no. 2, pp. 449–455.

Watson, Gary (1990): "Free Agency". In: *Free Will*, Gary Watson, ed. Oxford: Oxford University Press, pp. 96–110.

Williams, Bernard (1981): "Persons, Character, and Morality". In: *Moral Luck: Philosophical Papers.* Cambridge: Cambridge University Press, pp.1–19.

Williams, Bernard *Ethics and the Limits of Philosophy* (1986). Cambridge, MA: Harvard University Press.
Williams, Bernard *Shame and Necessity* (1993). Berkeley: University of California Press.

Christoph Horn

What is Kant's Precise Answer to the Question 'Why Be Moral'?

The question, 'Why be moral?' is typically raised in situations in which an agent already knows what would be the right thing to do but asks for an additional motivation why he (or she) should implement the morally correct option. This hesitation to realize what should be done might often be traced back to the fact that acting morally is – or at least can be – somewhat detrimental (or not sufficiently advantageous) to the agent under consideration. Therefore the most convincing type of answer might be given by pointing out in which sense acting morally is part of the well-considered interest of the agent himself. This can simply be done by referring to external enforcing mechanisms such as legal sanctions or social pressure or by pointing at divine punishment in this life or a future one. A second possible answer is less easily available: one might claim that, by acting morally, the agent (directly or indirectly) receives something advantageous or good for himself. Morality is beneficial to the person who actualizes it. This can be meant in the sense of a gratification or some social recognition resulting from it or in the sense of inner peace of mind or some sort of self-content. A more sophisticated type of answer is that provided by ancient eudaemonism: there, moral agency is characterized as profitable for the agent himself by reference to our 'true' (rational) nature which is fully developed as soon as we are in in the possession of cognitive and moral virtues. By practising morality we are, according to that view, actualizing our nature and are gaining the happiness that is inherently connected to self-perfection.

Prima facie, all answers considered above seem to be excluded within Kant's moral philosophy. For Kant prominently defends a purely deontological position: fulfilling the demands of morality does not lie within the self-interest of the agent, and it does not amount to realizing one's genuine nature. Morality must be observed at any cost, even in cases of the strongest limitations of someone's happiness. But is that a correct reading? Does this mean that Kant has nothing to answer to the 'Why be moral?' question? The aim of this paper is to show that Kant, even if in an indirect sense, is somewhat closer than expected to what I described as the response of ancient eudaemonism. In order to corroborate this unusual reading, we should first have to look closer at the *rational agency*-interpretation of Kant's moral philosophy that seems to provide a promising way out of the problem (I), then at moral intuitionism which also has an interesting response to offer (II), before I try give a third answer that is closely

linked with the Kantian background theory that underlies his deontological model. This turns out to be surprisingly akin to eudaemonism, an affinity which I will discuss in the end (IV).

I

The 'Why be moral?' question would find a simple and convincing answer with regard to Kant if the *rational agency*-interpretation (as defended by O'Neill 1989, Hill 1992, Herman 1993, Korsgaard 1996, Guyer 2000 and others) was correct. According to the *rational agency*-reading, Kant should not be seen as a pure formalist whose approach to ethics is primarily based on an abstract universalization procedure; the categorical imperative is not (or not primarily) an algorithmic test of the logical consistency of maxims. Such a procedure is at least not at the heart of his moral philosophy, but it is only of indirect importance. Instead, Kant takes the rational agency of an individual, i.e. his (or her) capacity of setting ends, to be the decisive intrinsic, even absolute value. Seen under this premise, the maxims of someone's will (*Wille*) must be, following Kant, apt to foster the good of rational and self-determined freedom. The purpose behind the CI-procedure is hence not to test formal consistency, but to look at the congruence of possible maxims with the basic good of rational freedom. Additionally, these interpreters understand Kant as deriving the axiological character of all other goods from the central value of rational autonomy. The basic good thus has the rank of a higher-order good; it has a 'value-conferring status'.[1] The foundation which underlies Kantian ethics would then be a teleological or axiological one. Hence, if Kant defended this basic idea and a theory of goods derived from it, then the widespread impression that we are faced, in his ethics, with a purely formal procedure would be wrong. A certain formalism of his ethics would then result from the fact that rational agency is interpreted as a higher-order value which establishes an evaluative criterion for lower goods.

In fact, Kant seems to say in several contexts that freedom – 'freedom' in the sense of an enabling condition of our rational agency – is the crucial value that should basically orient our conduct as a meta-norm. One impressive passage can be found in the lecture notes of Collins (Ak. 27:344, my translation)[2]:

1 See Korsgaard 1996: 122 and 128.

2 "Worauf beruht denn das principium aller Pflichten gegen sich selbst? Die Freyheit ist einestheils das Vermögen, welches allen übrigen unendliche Brauchbarkeit gibt. Sie ist der höchste Grad des Lebens. Sie ist die Eigenschaft, die eine nothwendige Bedingung ist, die allen Vollkommenheiten zum Grunde liegt. [...] Der innere Wert aber der Welt, das summum bonum, ist die

> Whereupon does the principle of all duties against oneself rest? Freedom is on the one hand the capacity which gives all the others infinite usefulness. It (i.e. freedom) is the highest degree of life. It is the attribute which is a necessary condition that underlies all perfections. [...] The inner value of the world, the summum bonum, is the freedom of the will (*Willkür*) which is not necessitated to act. Freedom is hence the inner value of the world.

At first glance, the passage sounds as if Kant wanted to say that freedom is a fundamental, first, and non-derived value, a value that indirectly constitutes the worth of all other goods. One can try to strengthen this reading by combining it with the idea of practical self-relatedness, i.e. by some sort of *oikeiôsis*-concept as we know it from ancient Stoicism. This is founded on the idea that our practical self-relatedness is inescapable: as agents, we have to accept the identity of an end-following individual that wants to be successful in what he (or she) considers as a good; ultimately, one wishes to realize the good life as one understands it. Our self-relatedness has relevant normative implications. According to this basic idea, nobody can reasonably reject the necessary preconditions of his autonomy and his successful agency, namely the enabling conditions of his agency.

While the rational agency interpretation is systematically attractive, it finds insufficient support in Kant's writings. At first glance, it is not impossible to interpret Kant's position in this sense; the question of certain teleological elements within Kantian moral philosophy is a complex one and cannot be easily answered. Christoph Bambauer in his highly commendable study *Deontologie und Teleologie in der kantischen Ethik* (2011) has therefore rightly pointed out the ambiguities and shortcomings of a simple classification of Kant as a deontologist or teleologist. If we could ascribe to Kant a substantial good as the currency of his moral judgments, say, e.g., the freedom as the capacity to set ends, then we would come to an interpretation which is highly welcome for at least three reasons: (a) This interpretation brings Kant closer to the old tradition of European moral philosophy, (b) it leaves behind mere deontology and opens room for a Kantian account of happiness, character, virtue, welfare, and the meaningfulness of life, (c) it allows us to describe the Kantian position in terms of a systematically attractive constructivism instead of a strong metaphysical theory (see the classical studies of O'Neill 1975 and 1989).

But there is strong evidence against this reading.[3] In the 'Analytic' of the second *Critique*, Kant devotes his second chapter to the question how the categori-

---

Freyheit nach Willkür, die nicht neceßitirt wird zu handeln. Die Freyheit ist also der innere Werth der Welt."

3 See also the objections raised by Patrick Kain 2006, 2009 and 2010.

cal imperative and the theory of goods and values should be related. He relativizes (as we will see) the Humean approach which assumes that only an empirical account of pleasure and pain can be the starting-point of moral philosophy. In one of the core passages, Kant disclaims every type of goods-based approach (*KpV* 5:58):

> If the notion of good is not to be derived from an antecedent practical law, but, on the contrary, is to serve as its foundation, it can only be the notion of something whose existence promises pleasure, and thus determines the causality of the subject to produce it, that is to say, determines the faculty of desire. Now, since it is impossible to discern a priori what idea will be accompanied with pleasure and what with pain, it will depend on experience alone to find out what is primarily good or evil.

According to Kant, it would be unacceptable to derive the moral law from an antecedent good. But this precisely is what the *rational agency*-interpretation does. By relativizing the Humean approach, Kant dismisses each model of moral philosophy based on goods. All of them must consider goods, he thinks, as something derived from our experience; but what is morally strictly commanded or strictly forbidden cannot be described in empirical terms because of its invariance and immediacy.

Furthermore, there is a strong textual argument against the rational agency reading: passages which go into the same direction as our quotation above are rare in Kant's writings (if we accept at all the lecture notes of Collins as a Kantian writing). The prevailing number of texts develops the idea that the moral law commands rational agents immediately (not mediated by a perspective of a good or of his practical identity) to do what is necessary.

On my view, there is nevertheless some truth in the rational agency reading. We should not neglect that Kant has a theory of goods and values, a theory which is partly Humean and partly anti-Humean. What Kant in fact does is to derive values, in the Humean part of his theory, from desire and, in the anti-Humean part, from the moral law. The first, Humean element of this theory describes the value perspective which Kant calls that of self-love, whereas the second one constitutes the morally adequate value perspective. While the first is that practised by everyone in daily life, the second one is that which is morally obligatory. We *should* deduce the value of all other goods not from our desires, but from the act of commandment as Kant claims in a famous text from the *Groundwork* (Ak. 4:436):

> For nothing has any worth except what the law assigns it. Now the legislation itself which assigns the worth of everything must for that very reason possess dignity, that is an unconditional incomparable worth; and the word respect alone supplies a becoming expression

> for the esteem which a rational being must have for it. Autonomy then is the basis of the dignity of human and of every rational nature.

As the text emphasizes, there is, on the one hand, the CI-procedure which is characterized by immediacy and strictness. On the other hand, there is some sort of value-generating process. The moral commandment constitutes, in a second step, values. In the case of the famous and much debated Kantian version of 'human dignity', an individual bestows value to himself by following the instructions of moral law. Both elements, the immediacy of the commanding reason and the derivative generation of values, are confirmed by the doctrine of the 'fact of reason' (*Faktum der Vernunft*) in the second *Critique*, as we will see in section III.

II

Kant rejects every sort of eudaemonism as an adequate foundation of moral normativity. He is therefore unable to describe ethics in terms of benefits and losses with regard to our happiness. In the *Critique of Practical Reason*, he unambiguously declares that "all material principles are inappropriate for the highest moral law" (Ak. 5:41). If we can reasonably conclude that the *rational agency*-reading is wrong, there is a second interpretative temptation: namely to see Kant close to moral intuitionism. What he basically seems to do is to derive all value from a first immediate insight we cannot sufficiently account for. This insight has the character of a non-derived, unprecedented event. The textual evidence for such a reading might be seen in sections 6 and 7 of the first chapter of the second *Critique* where Kant develops his famous 'argument from the gallows' and then explains it by the idea of a 'fact of reason'. On the basis of this passage, many readers (e.g. G. Prauss in his monograph from 1983) believe that what Kant is claiming is what one might call the 'immediacy thesis' regarding practical reason. Kant would then be a moral intuitionist (Ak. 5:30):

> Suppose someone asserts of his lustful appetite that, when the desired object and the opportunity are present, it is quite irresistible. [Ask him] – if a gallows were erected before the house where he finds this opportunity, in order that he should be hanged thereon immediately after the gratification of his lust, whether he could not then control his passion; we need not be long in doubt what he would reply. Ask him, however – if his sovereign ordered him, on pain of the same immediate execution, to bear false witness against an honourable man, whom the prince might wish to destroy under a plausible pretext, would he consider it possible in that case to overcome his love of life, however great it may be. He would perhaps not venture to affirm whether he would do so or not, but he must unhesitatingly admit that it is possible to do so. He judges, therefore, that he can

> do a certain thing because he is conscious that he ought, and he recognizes that he is free – a fact which but for the moral law he would never have known.

As the text says, every one of us possesses an immediate consciousness of his obligation to act morally, and everybody knows of his capacity to follow this obligation. This fact is characterized as inescapable and undeniable shortly after the last quote (Ak. 5:32):

> The fact just mentioned is undeniable. It is only necessary to analyse the judgement that men pass on the lawfulness of their actions, in order to find that, whatever inclination may say to the contrary, reason, incorruptible and self-constrained, always confronts the maxim of the will in any action with the pure will, that is, with itself, considering itself as a priori practical. Now this principle of morality, just on account of the universality of the legislation which makes it the formal supreme determining principle of the will, without regard to any subjective differences, is declared by the reason to be a law for all rational beings, in so far as they have a will, that is, a power to determine their causality by the conception of rules; and, therefore, so far as they are capable of acting according to principles, and consequently also according to practical a priori principles (for these alone have the necessity that reason requires in a principle).

As the second quotation makes clear, Kant should not be characterized as a moral intuitionist. The immediacy of moral law is a claim not about certain contents or material aspects, but concerns solely the formal method of an appropriate examination of given maxims. When discussing the categorical imperative in the *Groundwork* and the second *Critique*, he describes a certain procedure (or even several procedures) which allow(s) us to test maxims with regard to their universalizability. In a second sort of procedure he wants us to test the non-instrumentalization (or incomplete instrumentalization) of 'humanity' in our own person and in all others. Also the doctrine of the 'fact of reason' cannot be taken in the sense of an immediate insight based on specific moral sense and its feelings and emotions. Having seemed very close to a *moral sense* conception in his ethics of the 1760s, Kant very clearly rejects this idea in the classical moral writings of the 1780s. He is quite unambiguous on this point when discussing moral emotions. The only morally adequate emotion is that of respect or reverence (*Achtung*) which he describes as 'non-pathological', and this is conceived as an effect of the moral law within us. What Kant thereby emphasizes is that moral law enforces morality at the cost of our self-love or self-interest. Kant does not reconcile our desire for happiness with the requirements of morality – at least not immediately. Only in a second step, under the title of a 'highest good' (*höchstes Gut*), does he try to combine, in the *Critique of Practical Reason*, the perspective of deontological morality with the teleological view of what we seek as our true good. We will come to this in a moment.

Kant famously rejects intuitionism, more precisely 'intellectual intuition', in his epistemology (*KrV* B307) – and rightly so. Generally speaking, moral intuitionism – as defended by G.E. Moore, H.A. Prichard, W.D. Ross or Max Scheler – is systematically highly unattractive for several reasons. First, moral intuitionism is unable to explain and to resolve conflicts between divergent moral beliefs or intuitions. Concerning intuitive dissonances, intuitionists can, according to their theory, do nothing more than hint at further intuitions, but they cannot account for the greater or lesser degree of legitimacy of an intuition. Therefore all intuitions are, *prima facie*, to be equally justified. But this would include even very doubtful intuitions (e. g. of a racist or sexist type). Second, they cannot deal with a lack of moral sense. A person lacking moral intuitions seems to present an insurmountable challenge to them since this state must be rather *explained* as some sort of defect than *rejected* on an argumentative basis. Furthermore, intuitionists cannot account for the phenomenal difference between better (e. g. well-reflected) or worse (e. g. premature) moral intuitions. That is, they cannot explain how someone can, during his moral reflections, relinquish a first intuition in order to get to a second one. Practical deliberation would be something like waiting for the right intuition to come. But how then to deal with 'wrong intuitions' (i. e. those which someone sees now as mistaken while he held them earlier)? I think that a very strong additional argument against intuitionism can be developed from the philosophical practice of moral thought experiments: Many people come to different intuitive conclusions with regard to the relevant parameters which are used in thought experiments like the ticking time bomb scenario. Imagine a terrorist group threatening a city with a bomb while one member of this group who has been arrested might be tortured during interrogation in order to get relevant information from him. In such a case, parameters like the severity of the danger (e. g. if we are confronted with a devastating nuclear or a less destructive conventional bomb) or the degree of cruelty of the acts of torture (from less to more brutal forms) seem to play a role for our moral intuitions concerning a possible legitimacy of torture. As this suggests, intuitions are related to the relevant parameters involved in a given case of moral judgment, such as degrees of pain, destruction, abasement or loss of autonomy.

III

Instead of building his moral theory on a value-based foundation or on irreducible intuitions, Kant emphasizes that the moral law *enforces* our adequate behavior at the cost of our self-love or self-interest. But this enforcement, though conceived as an immediate act of reason, is simultaneously presented by Kant

as a procedural one and can be methodically reconstructed. This excludes an intuitionist reading of the categorical imperative. Moral action goes back to immediate and strict commandments of reason; but these have no content – they are the objects of a procedure-based insight.

At first glance, the immediacy claim sounds suspiciously similar to a *divine command*-ethics, and in fact, as many interpreters of Kant, from Hegel and Schopenhauer to Bernard Williams, have noticed, this resembles a theological normative account of conduct familiar from the Judaeo-Christian tradition. Note that Kant does not say that we should *follow* reason since it is our genuine nature; this is the ancient philosophical standpoint mentioned above. What he says is that we *must* obey reason without having any inclination to do so. If this impression were correct, what would that imply for the 'Why be moral?' question? Can we find anything in his writings that gives us a more elaborate answer to it? It seems as if, in this case, he could simply reject the question by indicating that commandments or instructions have to be fulfilled by those who are subject to the commanding instance.

Now we already saw that Kant believes that the act of commandment leads indirectly to the generation of values. For this reason, we can in fact identify a genuine Kantian answer. On a closer look, we see that Kant has a much more complicated story to tell, based on a rarely discussed theory of action which we can find especially in the *Critique of Practical Reason*. This action theory is surprisingly close to what we know from ancient or medieval accounts of the so-called 'human strife for happiness'. One might call this an *appetence theory* or *inclination theory*. Pre-modern (ancient and medieval) eudaemonism is usually embedded in a philosophical framework which can be described as a teleological action theory in the sense of an appetence theory. Some of its basic characteristics are the following ideas:

[1] Each action of an individual must always be directed towards an intended end or goal (as part of the definition of what it means to perform an action).
[2] By each intended end or goal the individual is unavoidably desiring or striving for a (real or at least a seeming) good.
[3] Ends or goals can be differentiated into instrumental (extrinsic) and intrinsic goods. The former are (normally or at least rationally) desired for the sake of the latter.
[4] Thereby we are faced in our rational agency with longer or shorter chains of instrumental goods; these are often interconnected and must always end up with intrinsic goods.
[5] Each action is part of a life-long continuum of such interconnected chains of goods.

[6] The continuum of goods finally amounts to one big comprehensive, final good.
[7] The one big comprehensive, final good of human life is happiness or flourishing life.

Systematically, the most provocative claims contained in appetence theories are theses [1] and [6]. We can call them the 'thesis of the purposiveness of each single action' and the 'thesis of the summation of all goods to a comprehensive final good'. Needless to say, Kant is far away from defending either of these points; nevertheless, he makes use of their theoretical foundations to a relevant extent in the second *Critique*. What is remarkable is that the entire passage on the fact of reason is strongly embedded in an action theory belonging to the tradition of such a teleological appetence theory. In order to make this plausible, we should first have a look at the terminology employed by Kant in this passage. It is strongly based on the concept of 'will' (*Wille*) and the somewhat enigmatic expression 'determination of will' (*Willensbestimmung*). Concerning his concept of *Wille*, it should be noted here that it is not yet as clearly distinguished from *Willkür* (arbitrary choice) as we know it, roughly ten years later, from the Introduction of the *Metaphysics of Morals*. Regarding the concept of *Willensbestimmung*, it is important to have in mind that the will is not the determining subject, but the determined object (i. e. in the grammatical sense of a *genitivus obiectivus*). Let us examine in more detail how Kant develops his version of an appetence theory in the second *Critique*.

As we already saw, Kant thinks that all previous moral theories are systematically mistaken (cf. *KpV* Ak. 5:39). All of them suffer from the same defect: they are based on a *material* determination of the will instead of a *formal* one. But what is a material determination of will as opposed to a formal one, and what makes the first inappropriate while only the second has to be seen as adequate? In the *Groundwork*, Kant defined the will as the capacity to conceive practical principles and to follow them (*GMS* 4:412). One might therefore interpret a formal determination of the will as leaving behind all the values, goods, and ends that characterize the material determination of the will. But if this reading were true, it would imply that Kant rejected the entire approach based on such key concepts as values, goods, and ends. But, apparently, he doesn't. As the notion of the will as an appetitive faculty already indicates, Kant operates on the basis of classical terminology. In the first chapter *Critique of Practical Reason*, we find the following difficult formulation (Ak. 5:22):

> All material practical rules place the determining principle of the will in the lower appetitive faculty (*im unteren Begehrungsvermögen*); and if there were no purely formal laws of

> the will adequate to determine it, then we could not admit any higher appetitive faculty (*kein oberes Begehrungsvermögen*) at all.

The distinction between a higher and a lower appetitive faculty is clearly part of traditional appetence theories. The quotation is about the constitution of a higher appetitive faculty; it is claimed that only a formal determination of will is able to constitute such a higher faculty. As the text clearly implies, Kant explains his position in classical terms of an appetence theory.

At this point, let us examine the concept of will as employed in the second *Critique*. Two of its basic features are these: each will, as conceived by Kant, must have a 'determining principle' (*Bestimmungsgrund*); and there are only two possibilities for a determining principle of the will: it must be either formal or material. In the case of a formal determination, reason is the determining factor; in the case of a material one, a feeling of pleasure and pain dominates the will. This can be seen from the following quotation (*KpV* 5:25):

> Reason, with its practical law, determines the will immediately, not by means of an intervening feeling of pleasure or pain, not even of pleasure in the law itself, and it is only because it can, as pure reason, be practical, that it is possible for it to be legislative.

In the case of a formal determination, the will remains in a 'pure' state since it is only influenced by the 'form of the law', not by any content, as Kant explicitly says a little later (Ak. 5:31):

> The will is thought as independent of empirical conditions, and, therefore, as pure will determined by the mere form of the law, and this principle of determination is regarded as the supreme condition of all maxims.

A pure will is in Kantian terminology not the same thing as a holy will (cf. *GMS* 4:414). Whereas a perfect or holy will needs no rectification since it is always directed towards the real good, the pure will is thought as a 'purified' one, rectified by moral obligation. This sheds some light on the idea that a will can be 'contaminated' by material objects and the feelings of pleasure and pain. All of these have, Kant believes, a deforming impact on the will.

Within an appetence theory, the will generally plays the role of a tendency or inclination which must receive an orientation by a determining principle. For Kant, there exist only two of them. Whereas a material principle aligns the will towards desire fulfillment and happiness, a formal determination directs it towards the real good. This implies that the will, according to Kant, is a goal-directed faculty; it must have an object or end. He explicitly affirms this in the *Critique of Practical Reason* (Ak. 5:34):

> Now, it is indeed undeniable that every will must have an object, and therefore a matter; but it does not follow that this is the determining principle and the condition of the maxim; for, if that is so, then this cannot be exhibited in a universally legislative form, since in that case the expectation of the existence of the object would be the determining cause of the choice, and the volition must presuppose the dependence of the faculty of desire on the existence of something; but this dependence can only be sought in empirical conditions and, therefore, can never furnish a foundation for a necessary and universal rule.

To sum up, important characteristics of the Kantian concept of will are the following: (a) The will is described as a 'capacity of conceiving rules', not as a faculty to choose; (b) it is understood as an 'appetitive faculty' (*Begehrungsvermögen*), not as a 'decision-making faculty' (*Entscheidungsvermögen*); (c) the will is always in pursuit of certain ends; (d) the will as such, being in a somewhat uncontaminated state, is conceived as 'pure will' and 'holy will' need not be rectified; it is always directed towards the (morally adequate) highest good (e. g. *KpV* V.32); (e) our will, by contrast, is an unstable one and hence must be an object of determination, as the expression 'determination of the will' (*Willensbestimmung*) indicates. Kant uses this term quite often having in mind that the will is the object of determination, not the subject. Given the fact that Kant distinguishes, as we saw, between a superior and an inferior appetitive faculty (*oberes/unteres Begehrungsvermögen*) and given these five features (a-e), it seems clear that 'will' in the more familiar sense of a capacity to make decisions cannot be meant by the Kantian concept of will. We are faced here with a typical pre-modern philosophical concept of will, that of a rational appetite, a concept famously attacked by Hobbes in the *Leviathan* (I.6) when he writes:

> The definition of the will, given commonly by the Schools, that it is a rational appetite, is not good. For if it were, then could there be no voluntary act against reason. For a voluntary act is that which proceedeth from the will, and no other. But if instead of a rational appetite, we shall say an appetite resulting from a precedent deliberation, then the definition is the same that I have given here. Will, therefore, is the last appetite in deliberating. And though we say in common discourse, a man had a will once to do a thing, that nevertheless he forbore to do; yet that is properly but an inclination, which makes no action voluntary; because the action depends not of it, but of the last inclination, or appetite. For if the intervenient appetites make any action voluntary, then by the same reason all intervenient aversions should make the same action involuntary; and so one and the same action should be both voluntary and involuntary.

The concept of will as a rational appetite rejected by Hobbes has a long history. We find it under the expression of *boulêsis* in Plato and Aristotle, under the terms *boulêsis* and *voluntas* in the Stoics and Neoplatonists, and under the notion of *voluntas* in medieval and early modern Scholasticism. It ultimately goes back

to a famous passage of Plato's *Gorgias* according to which willing (*boulesthai*) must always be directed towards rational and intrinsic goods. This is the reason, according to the Platonic Socrates, why rhetors and tyrants are powerless although they possess high influence in the cities (466d6-e2; translation T.H. Irwin, slightly modified):

> For I say, Polus, that both the rhetors and the tyrants have least powers in the cities, as I was saying just now; for they do practically nothing, I say, according to their will (*ouden gar poiousin hôn boulontai*), but do whatever they think is best.

Taken in itself, the will is a rational desire. By contrast with Kant, there is no contaminated will for Plato; he calls only an uncontaminated rational desire a will (*boulêsis*). A rational desire can strive only for a real good. Hence the will must be good as long as it is uncontaminated by false ends. This is what Kant says in the *Groundwork for the Metaphysics of Morals* (Ak. 4:414):

> A perfectly good will would therefore be equally subject to objective laws (viz., laws of good), but could not be conceived as obliged thereby to act lawfully, because of itself from its subjective constitution it can only be determined by the conception of good. Therefore no imperatives hold for the divine will, or in general for a holy will; ought is here out of place, because the volition is already of itself necessarily in unison with the law. Therefore imperatives are only formulae to express the relation of objective laws of all volition to the subjective imperfection of the will of this or that rational being, e. g., the human will.

One important observation that can be based on this is that the imperative character of Kantian ethics can ultimately be traced back to the idea of imperfect wills within an appetence theory. As Kant says according to *Moral Mrongovius II* (Ak. 29:606)[4]:

> The imperatives are taken from the idea of a perfect will and are valid as rules for my imperfect will; obligation is an idea of a perfect will as a norm for an imperfect one – God hence has no obligations.

In his *Metaphysics of Morals*, Kant explains what he means by the concept of *Willkür* by contrast to that of *Wille* as follows (Ak. 6:213):

> The faculty of desire (*Begehrungsvermögen*) may proceed in accordance with conceptions; and in so far as the principle thus determining it to action is found in the mind, and not in

4 "Die Imperative sind aus einer Idee eines vollkommenen Willens hergenommen und gelten als Regeln für meinen unvollkommenen Willen; die Pflicht ist eine Idee eines vollkommenen als eine Norm eines unvollkommenen Willens – Gott hat daher keine Pflichten".

> its object, it constitutes *a power of acting or not acting according to liking*. In so far as the activity is accompanied with the consciousness of the power of the action to produce the object, it forms an act of *choice* (*Willkür*); if this consciousness is not conjoined with it, the activity is called a *wish*. The faculty of desire, in so far as its inner principle of determination as the ground of its liking or predilection lies in the reason of the subject, constitutes the will (*Wille*). The will is therefore the faculty of active desire or appetency, viewed not so much in relation to the action – which is the relation of the act of choice – as rather in relation to the principle that determines the power of choice to the action. It has, in itself, properly no special principle of determination, but insofar as it may determine the voluntary act of choice, it is the practical reason itself.

The *Willkür* is the capacity of a conscious and arbitrary choice. The will, however, is an appetitive faculty determined by its goal. The *Wille* (in its perfect state) is practical reason itself. In his preparatory works for the *Metaphysics of Morals*, Kant adds (*Vorarbeiten zur Metaphysik der Sitten*, Ak. 23:248, my translation):[5]

> The human *Wille* must be distinguished from the *Willkür*. Only the latter can be called free and is directed only to phenomena, i.e. to actus which are determined in the sensible world. – For the *Wille* is not under the law, but is the legislator of the *Willkür* and is absolute practical spontaneity in the determination of the *Willkür*. This is why it is good in all human beings, and there is no unlawful willing. The maxims of the *Willkür*, however, since they are directed towards actions as phenomena in the sensible world, can be evil, and the *Willkür* as a natural capacity is free with respect to these laws (of the concept of obligation), by which it properly is not immediately determinable, but only through the maxims that can be according to it or against it.

In this text we are faced with a clear distinction between the *Willkür* as the faculty to make decisions in the empirical world, and the *Wille* which is not under the law, but absolutely spontaneous. The *Willkür* is 'free' in an empirical sense, and hence can make morally adequate or inadequate decisions. The *Wille* is 'free' in the sense of being the lawgiver and can hence not be unlawful. In the end, the *Wille* is more or less synonymous with the moral law.

---

5 "Der Wille des Menschen muss von der Willkür unterschieden werde. Nur die letztere kann frei genannt werden und geht bloß auf Erscheinungen, d.i. auf actus, die in der Sinnenwelt bestimmt sind. – Denn der Wille ist nicht unter dem Gesetz, sondern er ist selbst der Gesetzgeber für die Willkür und ist absolute praktische Spontaneität in Bestimmung der Willkür. Eben darum ist er auch in allen Menschen gut, und es gibt kein gesetzwidriges Wollen. Die Maximen der Willkür aber, weil sie auf Handlungen als Erscheinungen in der Sinnenwelt gehen, können böse sein, und die Willkür als Naturvermögen ist in Ansehung jener Gesetze (des Pflichtbegriffes) frei, durch die sie eigentlich nicht unmittelbar bestimmbar ist, sondern nur vermittelst der Maximen, sie jenem gemäß oder zuwider zu nehmen."

How does the end or goal of the will come into play? In his *Religion Within the Limits of Reason Alone*, Kant comes back to the question why even his formalistic moral philosophy presupposes the existence of a certain goal towards which the will must be directed. He gives the explanation that no will can be conceived as being without an end; in the case of the morally rectified will, the end of the will cannot be found as an antecedent one, but only as a consequence (Ak. 6: 4–5):

> But although for its own sake morality needs no representation of an end which must precede the determining of the will, it is quite possible that it is necessarily related to such an end, taken not as the ground but as the [sum of] inevitable consequences of maxims adopted as conformable to that end. For in the absence of all reference to an end no determination of the will can take place in man, since such determination cannot be followed by no effect whatever; and the representation of the effect must be capable of being accepted, not, indeed, as the basis for the determination of the will and as an end antecedently aimed at, but yet as an end conceived of as the result ensuing from the will's determination through the law (*finis in consequentiam veniens*). Without an end of this sort a will, envisaging to itself no definite goal for a contemplated act, either objective or subjective (which it has, or ought to have, in view), is indeed informed as to how it ought to act, but not whither, and so can achieve no satisfaction. It is true, therefore, that morality requires no end for right conduct; the law, which contains the formal condition of the use of freedom in general, suffices. Yet an end does arise out of morality; [...].

In this passage, Kant arrives at the following difficulty: How can an end that is by definition the antecedent object on which the will is directed (and hence must precede the appetence of the will), be seen as a consequence of the will after its rectification by the moral law? Generally speaking, the will must always be directed towards a preceding end and can never create this end out of the blue. Kant resolves this problem by appeal to the idea that the will can be seen as pure in an 'unfallen' state of human beings whereas now, under the conditions of its 'intelligible fall', it must first be re-directed in order to return to its original end. Morality must in this sense be prioritized to happiness. What Kant describes here is an inversion of the pre-modern procedure in ethics, but one that is still grounded on its appetence theories.

Already in the second *Critique*, we find a reflection on this necessary inversion of the procedure, namely under the title of a 'paradox of method' in the second chapter of the 'Analytic' (Ak. 5:62–3).

> This is the proper place to explain the paradox of method in a critique of practical reason, namely, that 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. In fact, even if we did not know that the principle of morality is a pure a priori law determining the

> will, yet, that we may not assume principles quite gratuitously, we must, at least at first, leave it undecided, whether the will has merely empirical principles of determination, or whether it has not also pure *a priori* principles; for it is contrary to all rules of philosophical method to assume as decided that which is the very point in question. 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.

What Kant describes here is precisely his idea of deontology: an adequate moral philosophy, he claims, should not derive the concepts of good and evil from the factual direction of the will, but inversely, first determine the will by the moral law (which gives us an immediate concept of good and evil) and then see to which end the will must be directed. All of this passage is described according to the letter of appetence theories, even if not according to their spirit. In the context of this impressive quotation, Kant starts with the Scholastic formula *nihil appetimus nisi sub ratione boni* (Ak. 5:59) and then explains how this sentence is to be correctly understood: namely by interpreting the underlying *bonum-malum* dichotomy not in terms of benefits and detriments (*Wohl und Weh*), but in terms of the moral good and the moral evil (*Gut und Böse*). This passage corroborates our observation that Kant preserves the theoretical framework although he radically revises its intention.

IV

We can now see what sort of answer Kant provides to the 'Why be moral?' question. It is true that his basic approach shows similarities with some sort of divine command ethics, namely in the sense that what is morally appropriate and what must be done by the agent should not be described as based on human (genuine) interest or his (true) nature. On the contrary, morality must be observed at the possible cost of all non-moral values, goods, desires, and the happiness of the agent. Moral commandments or instructions thus radically limit and relativize the desires or preferences coming from our self-love. But nevertheless, it would be false to assume that Kant sees no relationship between the moral commandments that have to be observed and someone's happiness. What he claims

is that our will (that is in fact contaminated by pathological desire) must be rectified by the imperative form of the moral law.

Here the basic structure of appetence theories comes in. The will is not conceived as a capacity to choose freely and arbitrarily, but as a capacity to strive for an end. Kant uses the notion *Wille* quite generally: it stands for a collective term indicating the direction of all desires and interests which can be ascribed to an agent. Taken in this sense, the Kantian concept of will expresses the basic goal-directedness of an individual.

Now Kant maintains that the human will in its normal or default state is always contaminated by material ends and hence always directed towards happiness in the sense of desire fulfillment. In this doctrine, Kant is, to a certain extent, a Humean. But when the moral law emerges and 'formally' determines the will, then the goal-directedness of the will is not simply interrupted or limited, but re-oriented. The will, i.e. the human striving for the real good, is then vectored towards the genuine value of human beings, the highest good – which is described, in the second *Critique*, as happiness in the sense of desire fulfillment according to one's moral dignity. As this consideration shows, it is not the case that only a material determination of the will leaves its goal-directedness intact; also the formal one preserves the purposiveness of the will. But after a formal re-orientation of our basic inclination, the will has a new, morally appropriate final end. The will has then become a pure one, even if not a holy one. Human nature is, in its starting-position, something bad or 'fallen'; it needs to be rectified or restituted by the rationality of the moral law.

Seen from this perspective, the moral law urges us into a direction where our true and 'unspoiled' nature is restituted. If this is correct, a Kantian answer to the 'Why be moral?' question emerges. This answer is based on an indirect value perspective. While we are rationally forced by the moral law (direct sense), we can additionally be interested in this enforcement (indirect sense) since the moral law ultimately leads us to the only appropriate sort of happiness: that which is qualified by our moral worth. By being forced to neglect our happiness, the moral law realigns us – and precisely thereby, it finally brings us to our true nature and happiness.

Bibliography

Bambauer, C. 2011: *Deontologie und Teleologie in der kantischen Ethik*, Freiburg/München: Alber.

Brandt, R. 1995: Kants "Paradoxon der Methode". (Ein Stenogramm), in: R.W. Puster (ed.), *Veritas filia temporis. Philosophiehistorie zwischen Wahrheit und Geschichte*, Berlin/New York: de Gruyter, 206 – 216.

Guyer, P. 2000: *Kant on Freedom, Law, and Happiness*, Cambridge/New York: Cambridge University Press.

Herman, B. 1993: *The Practice of Moral Judgment*, Cambridge, Mass.

Hill, Th. E. 1992: *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca, NY: Cornell University Press.

Horn, Ch. 2002: "Wille, Willensbestimmung, Begehrungsvermögen", in: O. Höffe (Hg.), *Kant, Kritik der praktischen Vernunft* (Klassiker Auslegen Bd. 26), Berlin: Akademie, 43 – 61.

Hudson, H.H. 1991: "*Wille*, *Willkür* and the Imputability of Immoral Actions", in: *Kant-Studien* 82, 179 – 196.

Irwin, T. 1996: "Kant's Criticism of Eudaemonism", in: S. Engstrom/J. Whiting, J. (eds.), *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty.* Cambridge: Cambridge University Press, 63 – 101.

Kain, P. 2006: "Constructivism, Intrinsic Normativity and the Motivational Analysis Argument", in: H. Klemme/M. Kühn/D. Schönecker (eds.), *Moralische Motivation. Kant und die Alternative* (Kant-Forschungen 16), Hamburg: Meiner, 59 – 78.

Kain, P. 2006: "Realism and Anti-Realism in Kant's Second *Critique*", in: *Philosophy Compass* 1, 449 – 465.

Kain, P. 2009: "Kant's Defense of Human Moral Status", in: *Journal of the History of Philosophy* 47, 59 – 102.

Kain, P. 2010: "Practical Cognition, Intuition, and the Fact of Reason", in: B. Bruxvoort Lipscomb/J. Krueger (eds.), *Kant's Moral Metaphysics: God, Freedom, and Immortality*, Berlin/New York: de Gruyter, 211 – 230.

Kant, I. 1902 ff.: *Gesammelte Schriften*. Herausgegeben von der Preußischen Akademie der Wissenschaften (Akademie-Ausgabe = Ak.). Berlin: de Gruyter.

Korsgaard, Ch.M. 1996: *Creating the Kingdom of Ends*, Cambridge/New York: Cambridge University Press.

O'Neill, O. 1975: *Acting on Principle: An Essay on Kantian Ethics*, Cambridge/New York: Cambridge University Press.

O'Neill, O. 1989: *Constructions of Reason: Explorations of Kant's Practical Philosophy.* Cambridge: Cambridge University Press

Prauss, G. 1983: *Kant über Freiheit als Autonomie*, Frankfurt a.M.: Klostermann.

Sherman, N. 1997: *Making a Necessity of Virtue: Aristotle and Kant on Virtue.* Cambridge: Cambridge University Press.

Stekeler-Weithofer, P. 1990: "Willkür und Wille bei Kant", in: *Kant-Studien* 81, 304 – 320.

Stratton-Lake, P. 2000: *Kant, Duty and Moral Worth*, London: Routledge.

Stratton-Lake, P. 2006: "Moral Motivation in Kant", in: Bird, G. (ed.) *The Blackwell Companion to Kant*, Oxford: Wiley-Blackwell, 322 – 334.

Iddo Landau

The "Why Be Moral?" Question and the Meaning of Life

1

Of the many possible ways of approaching the question "Why be moral?", I focus in this paper on one, which links the question to the notion of the meaning of life. Following the work of David Wiggins (1976, pp. 348–349), R. W. Hepburn (2000, p. 262), Kai Nielsen (2000, pp. 237, 242–250), Susan Wolf (1997a, pp. 208–213; 1997b, p. 304) Berit Brogaard and Barry Smith (2005, p. 443), and many others, I take the "meaning" in "the meaning of life" to have much to do with worth or value. Discussions of the meaning of life complain that life does not have sufficient worth, or wonder what might give life sufficient value, or celebrate the finding of something of sufficient importance in life, etc. (In what follows I will use the terms *worth*, *value*, and *importance* interchangeably.) Take, for example, Tolstoy's narration, in his semi-autobiographical *My Confession*, of how he came to feel that his life was not meaningful. At a certain point in his life he started looking at all he had achieved and asking himself "What of it?" and "So what?" He was the greatest Russian author: so what? He owned a lot of land: what of it? (Tolstoy 1983, pp. 26–27). I suggest that the "so what?" and "what of it?" questions exclaim that what seemed to him of sufficient worth stopped appearing so. What troubled Tolstoy was the feeling that all that has been described above was not in fact of sufficient value.

Similarly, Thomas Nagel (1986) argues that from the objective, broad perspective of the whole cosmos and time, *sub specie aeternitatis*, our lives are not as meaningful as we would like them to be. He mentions that, seen from that broad perspective, our influence on the world is negligible; if we had not lived, nothing much would have changed, in the long run, for the world at large. Put differently, seen from the broad perspective, our death and our life are inconsequential or unimportant. Moreover, our coming into existence is contingent: we could have easily not been born. Nagel's arguments suggest that, from the *sub specie aeternitatis* perspective, our lives do not seem to have much value; he is discussing the insufficient worth of human beings (when seen from that broad perspective).

Discussions I have had with people who thought that their lives were meaningless, or were searching for what would make them more meaningful, also confirm that those people were preoccupied with issues of worth and value in

their lives. They had not found something of value, or what had been worthwhile in their lives was taken away, or they no longer saw what could be of sufficient worth in their lives. A person who lost a beloved brother felt that life was meaningless because something very valuable he once had in his life was now gone. Another person said that she found life meaningless upon painful disappointment with a political movement for which she had sacrificed much; again, something that had endowed great value ceased to do so. And an able and ambitious biologist I knew confided that she felt that her life was meaningless because, in spite of her many efforts, she failed to reach what she considered to be the very top of her profession. For her, this was a sufficient reason to judge her life to be meaningless. What was of extreme value to her (whether for good reasons or bad) was the public or professional recognition of her peers that she was at the very top of her profession; when it became clear to her that she would not achieve that, she felt that her life was meaningless.

All other discussions of the meaning of life seem to involve similar preoccupations. Complaints that life is meaningless translate well to claims about the lack or insufficiency of aspects of value in that life. The search for meaning translates well to a quest for aspects of sufficient value. Therefore, I will henceforth treat the meaning of life as the value, or worth, in life. A meaningful life is one that has a sufficient number of aspects of sufficient value. A meaningless life is one without a sufficient number of aspects of sufficient value. (This is why people sometimes describe meaningless lives as "empty"; they are empty of sufficient value.) To make a meaningless life into a meaningful one, or to make an already meaningful life into a more meaningful one, we should increase what is of worth in our lives.

Now if one accepts what I have suggested here, I believe that we already have a beginning of a reply to the question "Why be moral?" Those who want to have meaningful lives have a reason to be moral, because being moral increases the value, or worth, in our lives. Morality makes our lives more meaningful.

2

However, there are many ways in which this suggestion may be criticized. One way would be to point out that some theories of the meaning of life are subjectivist, and as such imply that meaningful lives need not be moral at all. Under subjectivism, highly immoral lives could be highly meaningful since subjectivist theories do not rely on objective criteria but take the endorsement of beliefs, feelings, or sensations that one's life is meaningful to be a sufficient condition for leading a meaningful life. Richard Taylor (1970, p. 265), for example, argues

that "if Sisyphus had a keen and unappeasable desire to be doing just what he found himself doing, then ... it would ... have a meaning for him."[1] But this, of course, also implies that if a murderer has a keen and unappeasable desire to murder, his life, too, is meaningful. This is true, of course, not only on Richard Taylor's subjectivist theory of the meaning of life, but also on all other subjectivist ones. Since, for them, endorsing a certain belief, feeling, or sensation about one's life is a sufficient condition for leading a meaningful life, they allow that radically immoral lives could be meaningful.

Various considerations, however, suggest that subjectivist theories of meaningfulness are too problematic to accept. One consideration returns to the characterization of meaningfulness as value. As suggested above, we understand meaningful lives to be lives that include a sufficient number of aspects that are of sufficient worth or value. When the overall value passes a certain threshold, life becomes meaningful, and when it continues to increase, an already meaningful life can become even more meaningful. Meaningfulness, then, rests on value. But this is a reason to reject subjectivism as regards the meaning of life, because we commonly think that people can be wrong in their evaluations, including their self-evaluations. For example, one may believe oneself to be a good parent or spouse although one in fact is not. One may also wrongly think that one's scientific work is good when in fact it is not (and vice versa) or that one is a good pianist when one is not (and vice versa). We sometimes think that people are too strict with themselves and that they actually write better literature, or better philosophy, than they think they do, while at other times we believe that people have too positive a view of their achievements. But if we accept that one can be wrong in the evaluation of specific aspects of one's life, it is inconsistent to believe that one cannot be wrong in one's estimation of the overall worth of all the aspects. According to subjectivist understandings of the meaning of life, however, one cannot be wrong in one's estimation of the meaningfulness of one's life. If I feel or think that my life is meaningful, it is meaningful, and if I think or feel that it is meaningless, it is indeed meaningless.

A second consideration relies on the notion of reflective equilibrium, presented by Rawls (1971, pp. 48 – 51) and in wide use today, which proceeds by examining and revising our views, judgments, and intuitions by looking for their coherence with other views, judgments, and intuitions about similar and other issues, sometimes revising some of them for the sake of coherence with others. But subjectivist understandings of the meaning of life have some extremely

1 In a later paper, Taylor changes his view and mentions other, objective conditions such as autonomy, purpose, and creativity, but not morality (1987, 679 – 82).

counterintuitive implications, perhaps so counterintuitive that those who endorse the notion of reflective equilibrium (or want their theory to accord with very clear and strong intuitions for other reasons) have a reason to find subjectivism too problematic to accept. Subjectivist understandings of the meaning of life consider a person who, for example, thinks that his life is meaningful because he is of the same height as some tree on the Siberian plain (Taylor 1992, p. 36) or because he devours his own excrement (Wielenberg 2005, p. 22) to indeed have a meaningful life. This, however, would seem to many of us to be too odd to accept and too far from the regular use of the notion of a meaningful life.

A third, important consideration has been proposed by Charles Taylor (1992, pp. 31–41). Taylor points out that when we suggest that something makes our life meaningful, we do not mean that it does so because we just happen to think that it does. We think that that thing *really* makes life meaningful, and that imparting or arriving at meaningfulness is not arbitrary: it is not the case that anything else could have done so as well. We do not think that something is meaningful because we "just feel like it." This is so because the very idea of meaningfulness includes the notion of not being arbitrary. What is meaningful to us cannot be just anything whatsoever. In order to be meaningful, it has to have a certain quality or characteristic that is objectively meaningful, and that quality is what causes us to choose that particular thing. Even if we asked one of those mythical figures such as the excrement eater why they think that what they do is meaningful, they would most probably not answer "just so," or "just because I happen to be thinking about it now." They would give us a reason or tell us a story—perhaps a bad reason or a bad story, but a reason or a story just the same—to explain why what they do is really or objectively important. Perhaps the reasons or stories will have to do with some religious practice or ritual or involve an important symbol for something worthy and great. Or there may be some other reason or story, but we would very likely receive some reason that refers to what is taken to be "really" the case. Our informants are likely to tell us that what they do has to do with some objective worth, and that it is because of that objective worth that they endorse the activities that they do and that render their lives more meaningful. According to Taylor, then, our use of the notion of meaningfulness already presupposes objective rather than subjective worth.

I might add that the same is true of the notion of meaninglessness. People who say that their life is meaningless tend to distinguish very clearly between reporting a feeling or a thought, on the one hand, and making a claim about the way they believe their life really, objectively is, on the other hand. When they claim that their life is meaningless, they are not saying that they just feel awful, but are claiming that their feeling has to do with more than a feeling and that their life, or existence in general, is really, objectively, not of sufficient

value. Some may be unsure whether they are just in a bad mood and having dark thoughts or whether they are actually correctly conceiving the way existence is, but they distinguish very clearly between the two notions and take them to have different implications.

3

Another criticism of the claim about the relationship between morality and meaningfulness may accept that meaning in life has to do (at least in part, if not entirely) with objective conditions, while denying that these conditions have to do with morality. Indeed, some theories of the meaning of life that present both subjective and objective conditions for meaningfulness do not mention morality at all, thus allowing for highly immoral lives to be considered meaningful. A. J. Ayer, for example, posits "one's standing in one's society and the historical influence … that one exerts" as the objective conditions of meaningfulness (1990, p. 196; see also p. 194). For Paul Edwards (2000, pp. 143–144), a meaningful life is one in which one's actions relate to "some dominant, overall goal or goals which gave direction to a great many of the individual's actions" and in which one's attachments "are not too shallow." And for John Kekes the objective conditions of meaningfulness are successful rather than futile activities whose success relates to objective conditions in the natural world (2000, p. 32). But since, for Ayer, one's prominent standing in one's society need not have morally beneficial results, and one's historical influence need not be a morally positive one, his objective criteria allow for people who use their social power in their societies in quite horrid ways to have meaningful lives. Likewise, since Edwards does not hold that one's non-shallow attachments and overall goals must be moral, his objective criteria, too, allow for very evil individuals to be considered as leading meaningful lives. Edwards is aware of these implications and openly endorses them, accepting claims such as "as long as I was a convinced Nazi … my life had meaning … yet most of my actions were extremely harmful" (2000, p. 144). Similarly, since many people successfully realize quite immoral projects in the objective world, Kekes's criteria imply that if Jack the Ripper was successful in carrying out his plans, he too would have had a meaningful life. Kekes, too, is aware of the implications of his position, and writes "that immoral lives may be meaningful is shown by the countless dedicated Nazi and Communist mass murderers … [who] may be successfully engaged in their projects, derive great satisfaction from them, and find their lives as scourges of their literal or metaphorical gods very meaningful" (2000, p. 30).

Some other objectivist theories take morality to be a contributing factor that increases meaningfulness, but not a necessary condition for meaningfulness. For example, Laurence Thomas (2005, p. 405) argues that "on the one hand, it seems too strong to say that it is impossible for an immoral person to lead a meaningful life. On the other hand, we should like to think that a morally decent human being ... is ... more favored to lead a meaningful life than an immoral person is." Thomas's criterion, too, then, allows that an immoral life that fulfills some other conditions for meaningfulness could well be meaningful.

Theories that accept that immoral lives could be meaningful take meaningfulness and morality to be independent of each other. Such theories suggest, then, that we can describe a life as having achieved a certain degree of morality, and we can also describe a life as having achieved a certain degree of meaningfulness, but these descriptions neither imply nor exclude each other. Following Kekes's example, consider Bill, whose life was not meaningful until he joined the Ku Klux Klan. Before becoming a KKK member, Bill never believed in anything, never held a job for more than two days, and mostly moved, half drunk at best, from one bar to another. However, after he joined the Klan his life became more coherent and focused; it was now dedicated to an ideal and had a purpose (i.e., realizing some violent white-supremacist platform). He now had something to believe in, experienced self-worth and contentment, and had a considerable (murderous) effect on the lives of other people. Bill, the argument would go, indeed did not have a moral life; but he did have a meaningful one. We may take morality to be more important than meaningfulness and thus condemn Bill's life as immoral, even if meaningful, judging that it would have been preferable if he had not had this meaningful but immoral life but had instead remained an unfocused drunkard. Likewise, we may wish that he had had a less rather than more meaningful life, since then he would have been less effective and inflicted less harm. Still, the argument would go, Bill's life was meaningful. Just as radical immorality can be consistent with, say, good taste in music, a high IQ, or a thorough knowledge of classical literature, so too can it be consistent with meaningfulness. But if meaningfulness can be consistent with both morality and immorality, then the wish to have a meaningful life cannot function as a reason for being moral.

The view that lives such as Bill's can be objectively meaningful, however, conflicts with the common conception, mentioned at the beginning of the paper, of what a meaningful life is. As argued above, a meaningful life is a life that, overall, has a sufficiently high degree of worth or value. But if this is the case, life cannot be very low in, say, morality yet very high in meaningfulness. If a sufficient degree of worth is a necessary condition for meaningfulness, then morality and meaningfulness are not independent of each other, since mor-

ality, for better or for worse, affects the overall value of one's life and thus its meaningfulness. If we judge Bill's life to have a very low value overall because of its radical immorality, it cannot at the same time be a meaningful life.

This relates to another difficulty in Ayer's, Edwards's and Kekes's positions. According to them, we may take Bill's life to have become more meaningful once he joined the KKK and at the same time be *sorry* that his life became more meaningful. They take the claim that Bill had a meaningful life to be consistent with the claim that it would have been nicer if Bill had never existed at all. But the notion of a meaningful life, I suggest, is a laudatory, honorific notion that has positive connotations. A meaningful life is not a life that it is better not to have had; meaningfulness is a positive value that we want people to have and to increase. It is a concept that functions much like "heroism" or "wisdom." We may describe a certain SS soldier as bold, daring, or even brave. But we would not normally describe him as a *hero*, since for us "hero" has positive connotations. A hero is a person who behaves boldly and endangers himself for good causes. Likewise, we may describe a serial murderer or rapist who managed to evade the police for a long time as smart, clever, or intelligent, but we would not normally describe him as *wise*, a term we reserve for people who use their intelligence to gain understanding and knowledge that we see as constructive and helpful. Like heroism, wisdom, and some other terms (e. g., maturity), meaningfulness, too, is a laudatory term. Hence, it would be odd to suggest that people like Bill had had a meaningful life.

A third consideration that may lead us away from views such as Ayer's, Edwards's or Kekes's as regards the meaning of life and morality is somewhat tied to the previous ones: we take meaningful lives to be full of worth. We admire highly meaningful lives such as those of Mother Teresa, Bach, Martin Luther King, Shakespeare, Rubens, and Mahatma Gandhi, and we respect lives that are meaningful even when they have not reached such excellence. But our reaction to Bill's life is not one of admiration or respect but, rather, of abhorrence or contempt. The immoral behavior of rapists, blackmailers, thieves, liars, and thugs seems not simply wrong to us, but also despicable. We see such people as lowlifes and keep our distance from them not only because they make us angry, frightened, or cautious but because we are also disgusted by them. These are our reactions to what we find unworthy, or the opposite of worthy. It is for this reason that Thaddeus Metz (2002, pp. 805 – 807) has suggested that such lives include "antimatter," so to speak. They should be seen not only as lacking meaningfulness but, in analogy to negative numbers, as being on the negative part of the scale. Such a life is not only not meaningful; it is the opposite of meaningful: it is "anti-meaningful." But this suggests that im-

morality is inconsistent with meaningfulness. If we want to have a meaningful life, then, we have a reason not to act immorally.

4

It might be objected here, however, that being moral is not the only way of endowing life with worth or meaningfulness. We take not only Mother Teresa, Gandhi, and Martin Luther King, but also Einstein, Rembrandt, and Michelangelo to have had meaningful lives, although the meaningfulness of the latter had little or nothing to do with moral achievement. Rembrandt, for example, did not do much moral good, yet many would see his life as meaningful because of his artistic contribution. As Metz (2003, pp. 60 – 61) points out, creating artwork or making scientific advances can be meaningful activities even if they have no moral import. Although a meaningful life has to be evaluated positively, then, it need not be evaluated positively in terms of one's moral contribution or achievement. And this may suggest that replying to the question "Why be moral?" with "in order to have a meaningful life" may be problematic. In order to have a meaningful life you do not have to be moral. It may suffice that you be, for example, artistic or knowledgeable.

I agree. Although refraining from behaving in highly immoral ways is a necessary condition for having a meaningful life, behaving in highly positive moral ways is not a necessary condition for having a meaningful life. Hence Rembrandt's life could be seen as meaningful although he did not excel morally. But some minimal degree of moral behavior, or refraining from highly immoral behavior, is a necessary condition for meaningfulness. Once this condition is met, one's life can be deemed meaningful on the basis of value achieved also in other spheres of life. Hence, we should beware of claims such as "moral behavior is a necessary condition for a meaningful life," due to the ambiguity of "moral." We may take people to have behaved morally if they have committed no grave moral wrongs (when we say that such people lived morally we mean that they did not live immorally). But we may also take people to have behaved morally if they have helped others and performed deeds of charity or justice. Moral behavior is a necessary condition for a meaningful life only in the first sense, not the second.

But accepting that someone like Rembrandt could have had a meaningful life even though his contribution was not in the area of moral excellence does not undermine the suggestion that meaningfulness is an incentive for being moral. What has just been suggested is that if we want to have a meaningful life, we must refrain from highly immoral behavior. So meaningfulness gives

us a reason to be moral in the sense of seeing to it that we never become highly immoral. Moreover, it has also been argued above that if we want to have a meaningful life, we may follow various routes, one of which is excelling morally. So again meaningfulness gives us a reason to be moral, here in the sense of excelling morally. True, we may opt instead for other avenues to a meaningful life, such as the scholarly route, or the artistic route. Nevertheless, meaningfulness gives us a reason to be moral even if it gives us a reason to be other things as well. It gives us a reason to be moral as one option out of several. The wish to have a meaningful life arouses the motivation to develop in various possible directions, one of which is the moral direction. And this too is a reply, albeit a weaker type of reply, to the question "Why be moral?" If we want to have a meaningful life we *must* be moral in the sense of avoiding immorality, since this would undermine meaningfulness, and we *may* be moral in the sense of excelling morally, since this is one way of increasing meaningfulness.

However, I should qualify what I have just written. Although a meaningful life cannot include highly immoral behavior, it may include some immoral behavior; a meaningful life need not be impeccable. A generally worthy life can include, to some extent, behavior that we evaluate negatively, including behavior that we evaluate negatively from a moral point of view. Different kinds of behavior can balance each other out, to a degree, and we may deem a life that encompasses a limited degree of certain negative elements to be, overall, meaningful. Once a person crosses a certain threshold, however, we can no longer regard that life as having sufficient value and, therefore, as meaningful (of course, there will be some borderline cases). For example, we would probably continue to see Rembrandt's life as meaningful even if we learned that he had not always paid his debts on time or that there were some promises he had not kept. But we would not consider a Rembrandt who had to commit Jack-the-Ripper-style activities in order to find inspiration, or who sold his children into slavery in order to finance his artistic work, to have led a meaningful life.

5

Another possible objection to the link I have presented here between meaningfulness and morality has to do with cases in which, it seems, our wish to have meaningful lives or to increase the meaningfulness of our lives does not give us a reason to be moral but, on the contrary, gives us a reason to be immoral. Consider a case in which some immoral behavior, such as telling a small lie, committing a small theft, avoiding some responsibility, or failing to keep a promise or to return a debt, allows one to take advantage of a one-time opportunity

and thus, say, be accepted into an art school (or develop a beautiful love affair or receive an academic fellowship) that enables one to considerably increase the meaningfulness of one's life. This immoral behavior diminishes meaningfulness in one way, since it makes one's life slightly less worthy, it but also enables one to develop in other spheres or aspects of value in one's life (those having to do with, say, art, love, or scholarship) so much that one's overall life becomes much more meaningful. We can see, then, that meaningfulness sometimes does not answer the question "Why be moral?" but, rather, the opposite question: "Why be immoral?"

But I do not think that such examples undermine the claims made here. Even if meaningfulness, in some cases, gives us reason to be somewhat immoral, it also gives us reason to be moral because it always disallows any high degree of immorality. And while it gives us reasons not only to be moral but also to be, say, scholarly, or artistic, or loving, it always continues to give us a reason to be moral as well, as another option, or avenue, for making our lives meaningful. True, when meaningfulness gives us a reason to make our lives scholarly, artistic, or loving, it will in some specific constellations give us as well a reason to be slightly immoral. But that does not undermine the claim that, when "being moral" is understood as "refraining from being highly immoral," meaningfulness always gives us a reason to be moral, and when "being moral" is understood as "engaging in positive moral activities," meaningfulness gives us a reason to be moral in many, even if not in all, cases.

6

Another question may be whether we really gain any advantage when we employ the notion of the meaning of life to explain why we should be moral. I should note, first, that some people do not think that this question needs a reply at all. They see the requirement to be moral as self-evident, or an axiom, or just based on a very strong intuition, and they think that nothing more could or should be said about it. This seems to have been H. A. Pritchard's (1912) view in his famous paper that is commonly presented as having started off the debate. Those who believe that it is self-evident, or axiomatic, or strongly intuitive that we should be moral, and that nothing more could be said about it, will not think that anything has been gained in this paper, since nothing needs to be gained as regards the question "Why be moral?" to begin with. What has been argued up to now will be relevant only for those who think that it is sensible to ask this question and look for a reply. But even those in the latter group might argue that the reasons presented here for being moral do not really advance us; they just delay

the question. Assume that we should be moral because we want to have a meaningful life. But why should we want to have a meaningful life? If, again, we refer to self-evidence, or an axiom, or a very strong intuition, we have not progressed much.

However, I think that we have made some progress here. First, referring to the meaning of life, as we have done here, advances us since it shows in what sense we always need to be moral, in what sense we may be moral but may also opt for other ways of having a meaningful life, and in what cases we may also be slightly immoral. In other words, this paper does not merely suggest an axiom or intuition that might be more basic than that having to do with morality, but it also aims to specify the ways in which we should be moral (as far as our need to have a meaningful life is concerned). Yes, we should always be moral in the sense that we should never be highly immoral, but no, we do not always have to be moral in the sense of trying to achieve moral excellence. And it is also all right, in some restricted cases, to be slightly immoral.

Second, it seems that for many people having a meaningful life is more important than having a moral life. It is more self-evident and intuitive to such people that they should make sure that their lives are meaningful than it is that they should make sure that they are moral. For them, it will be profitable to begin with the notion of a meaningful life and proceed from there to the implications about morality.

Third, once we understand what makes life meaningful—namely that it be of worth or value—the reply to the question "Why have a meaningful life?" seems easier than the reply to the question "Why be moral?" To ask "Why have a meaningful life?" is to ask "Why have worth or value?" and the reply to that is that value is valuable, or that worth is worthy. Asking this question suggests that one wonders whether a tautology is correct, or that one does not understand what one is talking about, in a stronger way than that appearing when one asks "Why be moral?"

Fourth, morality seems to compete less successfully with other values, or with other inclinations we might have, than does meaningfulness. Many may well think that morality has some worth, but that this worth is in some cases overridden by the worth in some other values (and as shown in the examples above, it indeed sometimes is). Meaningfulness, however, as a supervening or second-order value, is not taken to be overridden in such ways. This, too, gives meaningfulness an advantage over morality as an intuitive starting point.

7

I have tried to present here a reply to the question "Why be moral?" But this reply is not meant to be exclusive. There may well be several valid reasons for being moral, and thus several replies to the question, just as there are several replies to the question "Why read books?" or "Why befriend people?" Perhaps some of the other replies will substantiate and argue for more demanding concepts of morality than I have done here, and some of them perhaps for less demanding ones. I suggest, however, that the reply "because it frequently enhances meaningfulness" is one helpful way of tackling this question.

Bibliography

Ayer, A. J. (1990): "The Meaning of Life." In: *The Meaning of Life and Other Essays*. London: Weidenfeld and Nicolson, pp. 178–197.

Brogaard, Berit and Smith, Barry (2005): "On Luck, Responsibility and the Meaning of Life." In: *Philosophical Papers* 34. No. 3, pp. 443–458.

Cottingham, John (2003): *On the Meaning of Life*. London: Routledge.

Edwards, Paul (2000): "The Meaning and Value of Life." In: Klemke, E. D. (Ed.): *The Meaning of Life*, 2nd ed. New York: Oxford University Press, pp. 133–152.

Hepburn, R. W. (2000): "Questions about the Meaning of Life." In: Klemke, E. D. (Ed.): *The Meaning of Life*, 2nd ed. New York: Oxford University Press, pp. 261–76.

Kekes, John (2000): "The Meaning of Life." In: French, Peter A. and Wettstein, Howard K. (Eds.): *Midwest Studies in Philosophy 24: Life and Death*. Malden, MA: Blackwell, pp. 17–34.

Metz, Thaddeus (2002): "Recent Work on the Meaning of Life." In: *Ethics* 112. No. 4, pp. 781–814.

Metz, Thaddeus (2003): "Utilitarianism and the Meaning of Life." In: *Utilitas* 15, pp. 70–50.

Nagel, Thomas (1986): "Birth, Death, and the Meaning of Life." In: *The View from Nowhere*. New York: Oxford University Press, pp. 208–231.

Nielsen, Kai (2000): "Linguistic Philosophy and the 'Meaning of Life.'" In: Klemke, E. D. (Ed.): *The Meaning of Life*, 2nd ed. New York: Oxford University Press, pp. 233–56.

Pritchard, H. A. (1912): "Does Moral Philosophy Rest on a Mistake?" In: *Mind* 21, pp. 21–37.

Rawls, John (1971): *A Theory of Justice*. Cambridge: Harvard University Press.

Taylor, Charles (1992): *The Ethics of Authenticity*. Cambridge: Harvard University Press.

Taylor, Richard (1970): "The Meaning of Life." In: *Good and Evil*. New York: Macmillan, pp. 256–268.

Taylor, Richard (1987): "Time and Life's Meaning." In: *Review of Metaphysics* 40, pp. 675–686.

Thomas, Laurence (2005): "Morality and a Meaningful Life." In: *Philosophical Papers* 34, pp. 405–427.

Tolstoy, Leo (1983): *Confession*. Trans. David Patterson. New York: Norton.

Wielenberg, Erik J. (2005): *Value and Virtue in a Godless Universe*. Cambridge: Cambridge University Press.
Wiggins, David (1976): "Truth, Invention, and the Meaning of Life." In: *Proceedings of the British Academy* 62, pp. 331–378.
Wolf, Susan (1997a): "Happiness and Meaning: Two Aspects of the Good Life." In: *Social Philosophy and Policy* 14, pp. 207–225.
Wolf, Susan (1997b): "Meaning and Morality." *Proceedings of the Aristotelian Society* 97, pp. 299–315.

Roe Fremstedal

Why Be Moral? A Kierkegaardian Approach

The present text focuses on what resources Kierkegaard offers for dealing with the question "Why be moral?" I sketch an approach to this question by presenting Kierkegaard's methodology, his negative arguments against the aesthete and the motive he offers for being moral. I conclude that Kierkegaard does provide motivation for assessing ourselves in moral terms, although his approach is more relevant to deontological ethics and virtue ethics than consequentialism.

Introductory Remarks on Methodology and Subject Matter

The fact that the question "Why be moral?" has been discussed many times suggests that the question is meaningful (cf. Hare 2002a, p. 95), even if a fully moral agent will hardly contemplate the question seriously. The present paper focuses on the relevance of the Danish philosopher Søren Kierkegaard (1813 – 55) for discussing this question. When I refer to "historical" issues, I do so mainly in order better to understand and illuminate the points that Kierkegaard tries to make that are still relevant to us. Thus, I am not concerned with historical issues as such but use them to enrich contemporary discussions. The important point for this paper is whether Kierkegaard's multifaceted ideas, or contemporary versions of them, can offer anything of interest to contemporary debates.[1] As a result, I have deliberately chosen to include Kierkegaard's pseudonymous authorship, particularly those parts of it which feature the so-called aesthetes and the ethicist, even though some Kierkegaard scholars go as far as maintaining that the pseudonymous writings cannot be attributed to Kierkegaard (cf. Poole 1997). The important thing for present purposes is what these writings bring to

1 My methodology will be what Gary Hatfield (2005, p. 91) has described as being "aware of the need for historical context to gain better access to past texts while still wanting to use those texts primarily as a source of raw material for solutions or answers to present philosophical problems". Thus, this paper belongs, at least mainly, to what is often referred to (especially by the Bennett generation) as analytic philosophy of history. A consequence of this is that I seek to use contemporary terminology rather than working with Kierkegaard's Danish and the Golden Age context (cf. Nadler 2005, p. 217). My main priority is to clarify Kierkegaard's claims, and to give them a charitable interpretation, although I will also say something about their strengths and weaknesses.

contemporary discussions, not whether they can ultimately be attributed to Kierkegaard.

The Argumentative Structure of *Either/Or*

By *aesthete* I understand an amoralist or someone who lives premorally since he is not fundamentally committed to morality. In what follows, I will focus on Kierkegaard's reflective aesthete (rather than the immediate or pre-reflective aesthete), since the reflective aesthete represents and concretizes moral skepticism.[2] The reflective aesthete only allows ethical considerations insofar as these considerations are subordinated to other concerns (and not given overriding authority). The aesthete does rely on prudential considerations,[3] but Kierkegaard sees these considerations as insufficient for morality proper since he works within the traditions of deontological ethics and virtue ethics (not utilitarianism).[4] Put in Kantian terms, the aesthete gives priority to empirical (material) principles over moral (formal) principles. The aesthete is not ruled by moral incentives but by competing incentives and principles. Much like Kant, Kierkegaard describes these competing incentives in terms of sensuousness, self-love, self-interest, and happiness (Knappe 2004, pp. 54 f., 94 – 97). The aesthete, then, is someone who is ruled by sensuousness, so that rationality and reflection serve sensuousness rather than morality. This intimate connection between the aesthete and sensuousness can be partially explained by the fact that Kierkegaard

2 Examples of immediate aesthetes include Don Giovanni in *Either/Or, Part I* and infants who are not yet capable of distinguishing between themselves and the surroundings. See SKS 2, 55 ff.; EO1, 47 ff.; SKS 17, 117 (Journal BB:25). I make use of the following standard abbreviations when referencing Kierkegaard:
CUP1 = *Concluding Unscientific Postscript to Philosophical Fragments*, Princeton: Princeton University Press 1992, vol. 1.
EO1 = *Either/Or, Part I*, Princeton: Princeton University Press 1987.
EO2 = *Either/Or, Part II*, Princeton: Princeton University Press 1990.
FT = *Fear and Trembling*, Princeton: Princeton University Press 1983.
SKS = *Søren Kierkegaards Skrifter*, vols. 1 – 28, Copenhagen: Gad 1997 – 2013.
SUD = *Sickness unto Death*, Princeton: Princeton University Press 1983.
UD = *Upbuilding Discourses in Various Spirits*, Princeton: Princeton University Press 2009.
WL = *Works of Love*, Princeton: Princeton University Press 1998.
3 See particularly "Rotation of Crops: A Venture in a Theory of Social Prudence" in *Either/Or, Part I* (SKS 2, 271 – 289; EO1, 281 – 300).
4 See Davenport 2001b, p. 273. Regarding virtue ethics, see Roberts 1998 and Rudd 2005, pp. 78 – 80, 99 – 105. Regarding deontology, see Knappe 2004, Chs. 3 – 5; Lübcke 1991, pp. 99 f.

takes aesthetic in the original Greek sense of *aisthesis*, as perception from the senses, although he associates the aesthetic with sensation, sensibility, and sensuousness (cf. Furtak 2005, p. 54; SKS 3, 29 f.; EO2, 21 f.).

Kierkegaard's *ethicist*, by contrast, is already fundamentally committed towards morality. He is convinced that he, or anyone who is seriously interested, is capable of being moral, since moral action relies on our volition rather than on luck or external conditions outside our control (Lübcke 1991, pp. 99 f.). *Either/Or* (and later pseudonymous works) can then be interpreted as a *dialogue* between various aesthetes (notably the pseudonym "A") and the ethicist (the pseudonym "Judge William" or "B" for short) that shed light on why we should be moral. Kierkegaard's pseudonymous writings develop the ethical position, or the ethical stage, by engaging in dialogue with other positions, notably the aesthetic stage. The different pseudonyms are used to describe different positions from within the first person perspective. It is just this dialogical approach that makes Kierkegaard's pseudonymous writings both interesting from a literary perspective and philosophically relevant for discussing the issue at hand. As Anthony Rudd has argued, *Either/Or* gives an extremely vivid literary portrayal of an amoralist instead of the colorless place-holder for a position of theoretical interest usually found in philosophical texts. Rudd elaborates:

> *Either/Or* as a whole challenges us to compare the self-portrait of the aesthete in Volume [Part] 1, with the description of him that emerges from Judge William's letters [to him] in Volume [Part] 2, and consider whether the Judge's account enables us to gain a better understanding of [the aesthete] "A" as he had appeared in his own writings. Within the work itself, the Judge challenges "A" to consider whether the ethical perspective will enable him to articulate more adequately what he already feels about his own life. (Rudd 2001, pp. 144 f.)

The ethicist Judge William is not just portraying the aesthetic and ethical forms of life but he argues against the aesthete.[5] The ethicist's first letter to A is called "The Esthetic Validity of Marriage" and his second letter is called "The Balance between the Esthetic and the Ethical in the Development of the Personality" (SKS 3, 13 – 151, 153 – 314; EO2, 3 – 154, 155 – 334). These two letters focus on the central importance of *love*, *selfhood*, and *freedom* not only for the aesthetic stage but also for the ethical stage. The ethicist argues that it is in the aesthete's true interest to become an ethicist, since the central notions of love, selfhood, and free-

5 I agree with Rudd (2012, p. 70) who says "I do think that Kierkegaard means to endorse Judge William's critique of the aesthetic stance, though he doesn't want to endorse all the Judge's positive views."

dom are better preserved ethically than aesthetically. Roughly, the idea is that without ethical commitment, love is episodic, lacking continuity and importance, while selfhood is unbalanced and freedom is negative, empty, and arbitrary.

For present purposes I will focus on the dialogue between the ethicist and aesthete, abstracting largely from the religious perspective that is also developed in Kierkegaard's authorship (including "Ultimatum" in *Either/Or, Part II*). Thus I will focus on what is traditionally referred to as the aesthetic and ethical stages, rather than the religious stage.[6] For the sake of argument this paper will accept the central point that mere prudence is insufficient for morality proper, since virtue and what is morally right cannot be reduced to a question of what brings happiness or well-being.[7] Kierkegaard even goes beyond this point by criticizing eudaimonism, and implicitly virtue ethics, for relying too much on prudential considerations.[8] By setting up the task in this way, that is, by disallowing arguments that are mainly prudential, utilitarian or even eudaimonistic, Kierkegaard makes it difficult to answer the question why we should be moral. This makes it even more interesting to see what, if anything, Kierkegaard can bring to the table.

Kierkegaard's intuition here might be sketched by saying that arguments which give us non-moral reasons or motives for being moral throw out the baby with the bath water, since we would then be moral for the wrong kind of reasons (something that would amount to legality instead of morality). Morality cannot be explained or justified in terms of anything more basic; it therefore needs to be (subjectively) recognized (Rudd 2012, p. 121).[9] On the other hand,

---

6 Unlike the ethicist, the religious person does not accept that we are capable of being morally perfect, but holds instead that morality presupposes divine grace. However, the religious writings generally presuppose the validity of ethics, arguing that philosophical ("first") ethics leads way to Christian ("second") ethics. Thus, the religious writings take philosophical ethics for given, much like revealed (transcendent) religion builds on natural (immanent) religion. Cf. Fremstedal 2014.

7 Recent scholarship on virtue ethics has argued convincingly that moral virtue is valuable in itself, not merely as a means for reaching happiness. See Annas 1993, pp. 125 – 128, 225 – 227, 260 – 290; Horn 1998, pp. 202 – 213, 220 – 224; Hare 2001, pp. 78 – 84.

8 Like Kant, Kierkegaard relies on arguments against eudaimonism that appear to have more force against hedonistic and Epicurean eudaimonism than Stoicism or even Aristotelianism (cf. Annas 1993, pp. 128, 236ff.).

9 Davenport (2001, pp. 79, 91) argues that Kierkegaard is a metaethical *internalist* in the sense that acting ethically means acting for the sake of the ethical, which means being motivated by the ethical rightness of the acts, rather than the goodness of their ends. Kierkegaard does not endorse the old saw of "the ends justify the means".

if we only give moral reasons or motives for being moral, then we stand in danger of begging the question, or so the aesthete might argue. Still, this does not necessarily prevent happiness or prudence from playing any role whatsoever; it only means that happiness and prudence must be a matter of secondary importance, while moral duty plays the primary role.[10] This means, arguably, that Kierkegaard's general approach to the issue at hand and ethics in general, is largely post-Kantian.

The ethicist sees the ethical task as the human task, arguing that the existential choice of oneself is identical to the choice of the ethical. Unless this account is to be circular, we must assume that there is some non-moral content to the self that a person should become (Evans 2006, p. 97). We need therefore to distinguish between the moral form of the self and its material content. The idea is that the aesthetic elements of the self are not to be eradicated but given a moral form. More specifically, sensuousness should not be eliminated but merely subsumed under morality.

The ethicist develops a quite sophisticated response to the aesthete A in *Either/Or, Part II*. Instead of merely condemning the aesthete on moral grounds, something that may appear moralistic and unhelpful, the ethicist sketches an *internal critique*; on the one hand, the aesthetic stage fails on its own terms, and on the other it is preserved in the ethical stage (Evans 2009, pp. 90 ff.; Ferreira 2009, p. 22). This argument can be said to involve a Hegelian *Aufhebung* of the aesthetic stage, where the aesthetic is partially negated because it is self-defeating, and partially recontextualized or lifted up to the ethical. Thus, apart from an external (transcending) critique of the aesthete on ethical grounds, the ethicist sketches an internal (immanent) critique that involves negative arguments as well as correctives. The ethicist argues on both aesthetical and ethical grounds, something that is also suggested by the title "The Balance between the Esthetic and the Ethical in the Development of the Personality". One example of this dual strategy is the claim that love needs moral obligations in order to endure; another is that the ethicist gains aesthetically by disciplining his desires. In what follows, I will focus on the internal critique of the aesthetic since it involves an indirect, dialogical, and maieutic approach that seems more effective and persuasive than a straightforward condemnation of the aesthete on moral grounds.

10 In this respect, Kierkegaard's approach overlaps with that of Kant, particularly the synthesis of morality and happiness found in Kant's doctrine of the highest good. Cf. Fremstedal 2014, Chs. 4 – 6.

The Argument from Despair

It seems that Kierkegaard's relevance to discussion of the question "Why be moral?" is reflected in his general methodology as much as in the content of his works. Still, this methodology is tied to the content of Kierkegaard's thinking. Kierkegaard can be said to develop a *via negativa* approach to ethics that claims that we only understand the ethical through its failure, through guilt, sin, and despair (Grøn 1997, p. 227). In German and Danish scholarship, this methodology is currently referred to as being "negativistic".[11] The methodology denies that we first have the ethical and then only afterwards have the possibility of failure. Rather, the normative task of being ethical, or becoming oneself, presupposes the possibility of failure, so that being ethical represents a problem (Grøn 1997, pp. 227, 261 f., 277). And the case of failure represents the rule rather than the exception insofar as ordinary human agents are concerned. In order to get a proper understanding of ethics, we therefore need to approach it indirectly by focusing on the aesthetic stage and how it can be said to involve despair. Kierkegaard's (Anti-Climacus') psychological analysis of despair can then be interpreted as disclosing ways in which one fails to be a moral agent, even though the ethical is inescapable.

In *Either/Or, Part II*, the ethicist develops a negative argument against the aesthete that I will refer to as the argument from despair. This argument tries to reduce the position of the aesthete *ad absurdum*. The absurdity, however, does not mainly take the form of a logical contradiction but rather involves a practical absurdity in the form of existential despair (and not merely something immoral). The central idea is that in order to avoid despair, one must transcend the aesthetic by choosing oneself, something that amounts to choosing the ethical. The ethicist thus offers a motive, rather than a proof, for transcending the aesthetic (Lübcke 1991, p. 97). This analysis of despair that is sketched in *Either/Or* is developed further in later works, notably *Concluding Unscientific Postscript*, "Purity of Heart,"[12] and *Sickness unto Death*.

The ethicist argues that the aesthetic view involves despair, either explicitly or implicitly. The part of *Either/Or* that describes the aesthetic stage from within, namely *Part I*, gives several indications of despair, particularly in the chapter "The Unhappiest One" (SKS 2, 211–223; EO1, 217–230). It is more difficult, how-

11 The main representatives are Michael Theunissen in Germany and Arne Grøn in Denmark. Cf. Theunissen 1991 and 1993; Grøn 1997.

12 The text commonly referred to as "Purity of Heart" is Part One of *Upbuilding Discourses in Various Spirits* (SKS 8, 115–250; UD, 3–154).

ever, to show that the different aesthetic views all imply despair. If the ethicist is to succeed in this, he has to distinguish between conscious and non-conscious despair and between authentic and inauthentic despair, as is done later in *Sickness unto Death*. The point then is that the *aesthete* (given the definition above) *has to be in despair*, even if he is not conscious or aware of it himself. This is the central claim that I will focus on in the following.[13] One relatively uncomplicated way to make sense of this claim is to say that despair is implicit in the aesthetic stage, and that it can be made explicit by reflecting upon it.

Despair is not a merely psychological concept that only refers to subjective experiences or a certain state of mind in which one lacks hope. In a way reminiscent of the concept of eudaimonia, which involves activity in accordance with virtue and (objective) well-being, despair involves objective, formal constraints that go beyond subjective experiences. In much the same way that it is possible to be unhappy even though one believes oneself to be happy, it is also possible to despair or lack hope (the expectancy of the good) without being conscious of it.[14] Inauthentic despair arguably implies self-deception or bad faith or that one lives in a way that is unstable (residing in a bubble) without being aware of it.[15]

Michelle Kosch (2006a, p. 154) argues that despair consists in an unwillingness to accept human agency (or selfhood) with all its particular conditions. On this interpretation, despair involves holding a false conception of oneself, a conception that does not reflect human agency and its conditions adequately. This interpretation allows for inauthentic despair, because it is perfectly possibly to have an inadequate conception of oneself without being aware of it. This line of interpretation makes it possible to explain, among other things, why despair involves self-deception and why despair consists in an act (in which one actively despairs) and not merely a psychological state. Despair is not just a result of suffering a loss, or experiencing hardship, but also something self-inflicted through guilt and sin (cf. Grøn 1997, pp. 143 – 153).

---

13 At this point I agree with Kosch and Rudd who have argued that there is a quite strong connection between the notion of despair in *Either/Or* and in later works such as *Sickness unto Death*. Kosch 2006a, pp. 142, 152 ff.; Rudd 2012, pp. 40, 70.

14 Kierkegaard appears to rely on a traditional Judeo-Christian understanding of the generic features of hope. To hope is to expect good (rather than just wishing for it). The object of hope must be possible to realize, yet uncertain; otherwise there would neither be room nor need for hope. In addition, what we hope for must be difficult to attain since there is hardly any need for hope if our goals are easily attainable (see Fremstedal 2014, Ch. 9).

15 SKS 11, 158 ff., cf. 136; SUD, 43 ff., cf. 20; Grøn 1997, pp. 125 – 140, 166 – 169.

This means, however, that human agency has a certain structure, a structure Kierkegaard scholarship has described mainly in terms of selfhood. This structure has important ramifications for the question of why one should take oneself to be a moral agent, subject to ethical demands. On this reading, the aesthete fails because his basic attitude towards his own existence involves a misconstrual of the nature of agency (Kosch 2006a, p. 143). Thus, the very structure of human agency or selfhood connects it to moral agency so that even aesthetic agency presupposes moral standards.[16]

In his influential analysis of human selfhood, Kierkegaard argues compellingly that selfhood is characterized both by freedom and necessity, transcendence and immanence. Neither of these two elements can be done away with; we cannot just identify with our given character (as Schopenhauer thinks we can) or with our freedom (as Sartre tends to think). However, these two elements always stand in a tense relation to each other, and we therefore tend to exaggerate either freedom or necessity. Still, it is only by reconciling freedom and necessity that we can become selves and overcome despair (Rudd 2012, pp. 48 f.).

Kosch (2006a, pp. 143, 149, 152 f.) offers a reconstruction in which the aesthetic stage collapses by denying the very possibility of choice, thus committing the aesthete towards passivity and fatalism. On this reading, the aesthete sees his own identity or character as essentially given, with no room for changes or modifications. At this point, Kosch invokes the systematic analysis found in *Sickness unto Death*, particularly the analysis of the "despair of necessity" which consists of a lack of possibility or freedom.[17] One clear problem with this interpretation is that it does not account for all the types of aesthetes Kierkegaard portrays. Although it shows that one attempt to escape the ethical fails, it does not preclude the possibility of other successful strategies.

One particularly important strategy that Kierkegaard devotes much attention to is the "despair of possibility" which consists of lacking necessity or limitations.[18] This type of position collapses by over-emphasizing freedom and self-creation, not by denying it as the fatalist does. Instead of seeing limitations as

16 Similarly, Theunissen (1991 and 1993) argues that the very notion of selfhood contains demands or normative requirements, at least implicitly.

17 The very similar "*despair of finitude*" consists of a lack of infinity (transcendence). Both these types of despair consist of believing that one is not capable of transcending facticity, or that one is not capable of breaking with the past. See SKS 11, 153 – 157, 149 – 151; SUD, 37 – 42, 33 – 35.

18 The very similar "*despair of infinity*" consists of lacking finitude. Both these types of despair imply that one wants to create oneself, without ethical restrictions, in order to get rid of the constraints of the present situation. See SKS 11, 151 – 153, 146 – 148; SUD, 35 – 37, 30 – 33.

something that makes real freedom possible, limitations are seen as a mere hindrance to freedom. This type of despair absolutizes freedom, understood negatively as the freedom from limitations (facticity). However, the result is that freedom itself is abstract and empty, since it does not allow for positive freedom to realize anything concrete. Neither does it allow for criteria making it possible to choose between different possibilities or alternatives, which means that it ends up with arbitrariness (and whims) because all possibilities are equally valid and equally abstract and empty.[19] This type of despair implies that one wants to create oneself, without ethical restrictions, in order to get rid of the constraints of the present situation. Kierkegaard argues that this implies not wanting to be the specific person one is, or not wanting to be positively free.[20] As a result the agent is double-minded or in despair, since he is split between necessity and possibility, immanence and transcendence. This makes sense if we keep in mind that one's possibilities only reside within the specific individual one is and in the particular situation one finds oneself in. And these possibilities are not morally neutral.

The ethicist argues that the solution to the problems posed by this type of despair lies in getting continuity or coherence in one's existence by appropriating necessity (facticity). He stresses that one's history is not solely a product of one's own free acts, but something closely related to the history of mankind as a whole (SKS 3, 171; EO2, 175). Hence, one's life can only have continuity if one sees that one stands in relation to other human beings, both the living and the dead (SKS 3, 239; EO2, 250 f.). When one sees reality as something one has appropriated, one sees oneself and one's surroundings in a *historical* and *social* perspective. In this context the ethicist stresses that the self is socially mediated: "[T]he

19 Scandinavians (and Germans) sometimes use the expression "like gyldig og likegyldig" here, something that means that the options are equally valid *and* indifferent.

20 Cf. Theunissen 1991, pp. 38 – 51; Grøn 1997, pp. 119 – 132, 182 – 189. Put in Kantian terms, we might say that the aesthete tries to absolutize the power of choice (*liberum arbitrium*; *Willkür*) and to do without pure practical reason (*Wille*). He thereby denies that negative freedom where one is free from alien causes involves positive freedom to be moral and autonomous. The ethicist, by contrast, has been taken by Kosch to endorse Kant's reciprocity thesis. The reciprocity thesis claims that negative freedom where one is free from alien causes involves positive freedom to be moral and autonomous, so that rational self-determination and transcendental freedom entail one another reciprocally. Following Schelling, Kosch argues, however, that this thesis is problematic: when freedom is understood as the capability for autonomy (self-determination), what is lost is freedom understood as the choice between good and evil. The result being, arguably, that moral evil is neither intelligible nor imputable. Thus interpreted, the reciprocity thesis implies that one is either moral or amoral. See Kosch 2006a, pp. 65 – 67, 87 f., 121, 129, 151, 169 – 173, 179. See also Allison 1995, Chs. 7 – 8.

self that is the objective [...] is a concrete self in living interaction with these specific surroundings, the life conditions, this order of things. The self that is the objective is not only a personal self but a social, a civic [*borgerligt*] self." (SKS 3, 250; EO2, 262) The upshot is that one cannot become a self, or synthetize possibility and necessity, without choosing the ethical (cf. SKS 3, 243 f., 249 f., 261; EO2, 255 f., 262 f., 274 f.). Kierkegaard's view, then, is that without choosing the ethical one either lapses into an unbalanced stress on restrictions and givenness or an equally unbalanced stress on freedom and voluntarism (Rudd 2012, pp. 104, 70).[21] This is also in line with the famous analysis of despair in *Sickness unto Death*, according to which inauthentic despair takes two basic forms, namely despair of necessity and despair of possibility, respectively (SKS 11, 145 – 157; SUD, 29 – 42).

Rudd's Reconstruction of the Argument from Despair

Recently, Anthony Rudd has attempted to reconstruct the argument from despair, arguing that Judge William's ethaical views are defensible and relevant to contemporary debates about morality.[22] Rudd summarizes his reconstruction of Kierkegaard as follows:

1. One can only avoid the necessity of judging one's life in moral terms by evading long-term commitments.
2. But to live such a life is to be in despair; for a life without commitments is one without purpose, and hence is one that makes it impossible to develop a coherent personal identity. (Rudd 2005, p. 69)

Rudd argues that a meaningful and fulfilled life requires a stable sense of self, something that

> can only be achieved through commitment to social roles and relationships which carry with them objective standards of assessment. One must become a participant in communities and the traditions which define them, and must develop the virtues necessary for such participation. The failure to do this will render one's life quite literally pointless. Without

21 See the previous page for a discussion of the despair of necessity (including references to Kosch 2006a).

22 Like Davenport (2012), Rudd (2012) is particularly concerned with the contemporary debate over whether non-moral caring involves implicit rational commitment to ethical values.

> any unifying *telos*, one's life collapses into a series of disconnected moments, and to live in this way [...] is to live in despair. (Rudd 2001, p. 139)

On this Kierkegaardian view, personal identity or selfhood is not something simply given but rather something that must be achieved through purposive moral action which synthetizes freedom and necessity, transcendence and immanence.[23] Rudd follows Bernard Williams in arguing that ground projects give meaning to life and continuity to our characters.[24] Ground projects are necessary if we are to develop a coherent personal identity (Rudd 2005, pp. 92f.). The crucial idea, however, is that *any project significant enough to give life purpose and meaning involves social interaction, practices, and institutions*.[25] However, these social practices and institutions always come with standards of assessment that are not only intersubjective, non-instrumental, and non-arbitrary, but also

23 Cf. Rudd 2005, p. 84. Rudd (and Davenport) argues that both actions and personal identity involve a narrative structure. Actions that are intelligible are purposive, involving (at least ideally) a decision, an act, and the attainment of a goal (Rudd 2005, pp. 84f.). Personal identity on the other hand requires not just single actions but also projects consisting of a pattern of purposive action. And it is only when our actions and identity belong to a larger narrative that they are intelligible and meaningful (cf. Davenport 2012). Rudd and Davenport thus connect moral agency to the narrative ideal, something that is controversial. John Kekes (2013) has recently formulated a criticism of the narrative ideal (as put forward by Alasdair MacIntyre) that makes the case that narratives are not necessary for a meaningful life. However, Rudd (2012) and Davenport (2012) have both responded to various objections against the narrative ideal, including objections developed by Kekes in his earlier publications (see especially Davenport 2012, p. 197). For the present purposes it seems unnecessary, and perhaps unfeasible, to discuss the narrative ideal thoroughly. However, it could be mentioned that Kekes' criticism concerns meaning in life rather than why we should be moral and that Kekes targets MacIntyre rather than Rudd, Davenport, or Kierkegaard. Kekes (2013, p. 71) argues that only an elite would be able to live according to the narrative ideal, whereas Rudd seem to hold that purposive action and participation in moral practices suffices for basic meaning in life. Kekes (2013) sees narratives as contingent human constructions, something Rudd (2012) and Davenport (2012) seem to deny by connecting narratives to objective meaning and moral realism.

24 Rudd 2005, p. 86; Davenport 2001b, p. 290. At this point, Rudd (2012, pp. 44f.) also makes use of Frankfurt's notion of "final ends" that one cares about for their own sake.

25 Rudd 2005, p. 94; cf. Davenport 2001b, p. 290. Rudd uses MacIntyre's definition of practice here. A slightly different approach is represented by Hare (2001, pp. 42–46) who argues for the necessity of assuming that what other people evaluate as good to pursue is at least roughly consistent with what I evaluate as good to pursue, since many of the goods I am likely to pursue depend for their achievement on the cooperation of others.

moral. Thus, significant projects involve sustaining non-instrumental personal relationships that require recognition of authoritative moral norms and ideals.[26]

A similar point is made by Rick Furtak (2005, p. 76) who argues that to "accept the roles of husband, judge, and friend (or mother, author, and confidante) is to accept certain beliefs about what is of value." Social roles and relationships involve intersubjective standards of behavior that are not merely dependent on my will, emotions, or subjectivity. Without such moral standards of assessment, Rudd argues that I would lack something that makes it possible for me to assess whether significant actions and projects are better or worse (Rudd 2005, pp. 71 f. and 2012, p. 110). Rudd (2012, p. 91) proceeds by arguing that there are good reasons for endorsing Harry Frankfurt's view that full selfhood requires a capacity for evaluation of my desires, dispositions, cares, and loves. However, this need for evaluation also involves an attempt to get things right (or get closer to being right); as evaluative beings, we cannot suppose that our evaluative judgments are incapable of being objectively correct or better (Rudd 2012, p. 95). We can only shape our identity as part of a rational process if we are able to make ourselves better or worse, judged by standards independent of our will. Rudd therefore concludes that "I have to ask, 'Do *I* consider this, or that good?' And this is why I think that the idea of *the* Good is unavoidable, if only as a regulative ideal. It is what my moral deliberation has to be constantly moving towards" (Rudd 2012, p. 141). Rational agency presupposes the possibility of rational examination of our higher-order cares and commitments in light of the idea of something that is objectively good (or at least better or worse). Without this possibility, the irrationality (or rather arationality) of our cares and commitments would cascade down the levels, and we would have no basis for thinking of ourselves as more than instrumentally rational agents.[27]

---

26 Rudd 2005, pp. 95, 115; Davenport 2001b, pp. 283, 297. Rudd 2001 prefers objective over intersubjective. Rudd (2012, Chs. 4 – 6) defends robust moral realism, and Davenport (2012, pp. 121 ff. and 2001b, p. 287) also defends moral realism and cognitivism.

27 Rudd 2012, p. 112. Rudd (2012, p. 115) concludes that "Rawlsian liberalism collapses into Schlegelian (or Rortian?) ironism – the valuing, not of *rational* choice, but of choice itself. But it is hard to see how such ironism can avoid collapsing into full-blown nihilism; for why should we treat the sheer power of choice as valuable, if there is nothing else that is genuinely valuable that it enables us to choose." Much like Rudd, Davenport (2001b, pp. 297 – 299) argues that moral standards provide a firm point outside of our first-order states that is much needed, since without such an objective basis, we have no stable ground for working upon ourselves; any attempt to better oneself will then be at the mercy of the contingencies of time. On this view, moral norms and ideals provide an *Anstoß* (in the Fichtean sense) by representing something radically different from subjective perspectives and first-order states (Davenport 2001b, pp. 297 f.).

Rudd's reconstruction of Judge William's argument for the ethical relies on moral virtues without the traditional idea of life having a final end (eudaimonia) that all human beings share.[28] Rudd argues that ground projects involve developing and exercising moral virtues.[29] He follows Peter Geach (a Thomist virtue ethicist) in arguing that "[w]e need prudence and practical wisdom for any large-scale planning [...], we need temperance in order not to be deflected from our long-term and large-scale goals by seeking short-term satisfactions. And we need courage in order to persevere in face of setbacks, weariness, difficulties and dangers."[30] The argument can be summarized as follows:

> Whatever projects one undertakes, one will need the virtues of courage, self-control, and practical wisdom, and also the virtue of honest perception [of oneself] [...]. In so far as one is committed to living in a society [...], one will also need the virtues of justice and benevolence, in some measure anyway. [...] the ethical task of developing the virtues is the same for everybody [...]. The need to cultivate the virtues derives from the need to engage in projects, and this derives from the need to live a coherent and meaningful life.[31]

On this view, there is a very close connection between the objectivity of moral values and the idea of meaning in life (Rudd 2012, p. 149). More specifically, a

28 Rudd 2005, pp. 78–80, 99–105; Kosch 2006a, pp. 146f. See also Davenport 2001b, pp. 291–293 who argues for a minimalistic telos consisting of self-integration and an existential unification of life-narrative (rather than eudaimonia). Against broadly eudaimonistic reconstructions of Kierkegaard like that of Rudd, Kosch objects unconvincingly that (1) agents who succeed according to aesthetic criteria are not only happy but also in despair, and (2) that Judge William does not dismiss aesthetic satisfaction but tries to preserve it in the ethical stage. However, for Kierkegaard the latter (2) seems to involve an *Aufhebung* of aesthetic values where they go from having absolute priority to being conditioned on morality. This way, pleasure can be consistent with moral duties or Kantian-Hegelian ethics. And the former point (1) suggests that a lucky aesthete can feel happy and therefore avoid psychological despair but this is perfectly compatible with inauthentic despair. This point is based on a questionable translation where "*lykkelige*" is translated as "indeed happy," giving the impression that the aesthetes are truly happy (SKS 3, 186; EO2, 192).

29 Rudd (2005, p. 94) follows B's distinction between personal virtues (courage, valor, temperance, and moderation) that are necessary for self-development and civic virtues (notably justice) that are necessary for participation in social life.

30 Peter Geach, *The Virtues*, p. 16 quoted in Rudd 2005, pp. 100f. Rudd (2005, p. 101) follows Iris Murdoch in stressing the importance of honesty with oneself.

31 Rudd 2005, pp. 108f., cf. 87. Virtue is described as "a disposition [...] giv[ing] constancy and stability to my character" (p. 108). Rudd does not think that Geach or classical virtue ethics succeeds completely in justifying the virtue of justice. Even though justice is necessary in order to secure cooperation and mutual trust among men, this hardly explains why I need to be just (pp. 101f.). Rudd concludes that justice remains problematic within the ethical stage, but not within the religious stage (p. 115), something that seems questionable at least exegetically.

coherent and meaningful life requires significant projects, something that involves social interactions and practices that presuppose moral norms. The aesthetic stage necessarily involves despair (or ennui) in the sense of lacking point and purpose with one's life. Since the aesthete does not want to commit to any projects, his life is pointless and without purpose.[32] And without anything that gives meaning to his life, the aesthete lacks something to unify the different parts of his life, something that makes it into a coherent whole with a clear personal identity. As a result, his life is nothing but a mere series of moments or episodes without a unifying structure. His life is ruled by a multiplicity of moods and situations, unlike the ethicist who relies on the unifying power of personality (Rudd 2005, pp. 75, 79 and 2001, pp. 138 f.).

An important part of this is the fact that Kierkegaard's aesthetes do not identify with social roles and commitments. For this reason, *Either/Or* describes the aesthetes as refraining from promises and obligations, and as warning against entering into friendship, marriage, and the acceptance of official positions (SKS 2, 284 – 287, 356; EO1, 295 – 298, 367). This, however, indicates a certain respect for the ethical (SKS 2, 356; EO1, 367). The idea is that one must avoid getting seriously involved with others; one must avoid commitment if one is to live aesthetically; otherwise, one will be trapped into social morality. One must therefore be able to avoid relationships, or to break them off by a sheer act of will (SKS 2, 286; EO1, 297; Rudd 2005, p. 71). Associated with this is the aesthete's view that morality is strict, harsh, boring, and rigid (Danish, *kantet*[33]) (SKS 2, 145, 356; EO1, 145, 367), since moral duties are opposed to our inclinations (SKS 3, 144; EO2, 146). It is not coincidental that this view resembles Schillerian criticism of Kant-

32 The aesthete writes that "My life is utterly meaningless. When I consider its various epochs, my life is like the word *Schnur* in the dictionary, which first of all means a string, and second a daughter-in-law. All that is lacking is that in the third place the word *Schnur* means a camel, in the fourth a whisk broom." (SKS 2, 45; EO1, 36)

33 The Hongs translate the Danish word "kantet" as "rigid" here. However, in this context the word "kantet" might be interpreted as an allusion to Kant or as a play on the word "kant". The Danish word "kantet" is based on the root "kant" and has the meaning rigid, edgy, rigorous, and inflexible (especially when "kantet" is used as a short for "firkantet"). The very meaning of the words "kantet" and "kant" in the Scandinavian languages fit perfectly the view that Kantian philosophy is overly rigid and rigorist. Moreover, the Danish reception of Kant closely associated Kantian ethics with moral rigorism (see Thuborg 1951, pp. 111 – 120). To this very day, Scandinavian philosophers can say that a philosopher is too "kantet," suggesting that he is too Kantian in the sense of being overly rigorous and inflexible.

ian ethics, since Kierkegaard's aesthete is heavily influenced by German Romanticism.[34]

One might worry, however, that this reconstruction leaves room for aesthetes who have infinite passion for non-moral projects without recognizing the validity of moral standards.[35] Why cannot a self-seeking egoist create his own projects or values, without caring about morality? Furtak (2005, p. 105) argues that when one loves nothing unselfishly, one must also "suffer the unbearable emptiness of a life in which there are no final ends, because nothing is cared about for its own sake." Kierkegaard goes a step further by arguing that the attempt to create one's own values, without caring about anything for its own sake, leads not only to values that are revocable and unstable but also to motiveless and arbitrary choice. When subject and lawgiver are identical, the subject influences the lawgiver; if one can bind oneself at will, one can also unbind oneself at will. This makes it possible to constantly change one's mind about what to do, to lazily concoct new tasks instead of realizing tasks that are given (SKS 8, 389 f.; UD, 294 f.). When values are mere contingent constructs of individuals who are fallible, imperfect, and non-moral, this leads to lawlessness or arbitrary experimentation not only as a possibility but also as a likely result. As long as it is not grounded in intrinsic values or objective norms, human freedom therefore threatens to collapse into an arbitrary and motiveless choice. Finally, Kierkegaard argues that consistency or wholeheartedness cannot be achieved without unconditional moral dedication. Although he concedes that an aesthete is capable of developing goals that may involve some unity and coherence (SKS, 3, 178; EO2, 183), he nevertheless insists that there is something superficial or inconsistent about such an aesthete (Davenport 2001b, pp. 299 f.). This is a claim to which we will turn in the next section.

Despair as Double-Mindedness

Rudd abstracts from much of what Kierkegaard says about despair. For Kierkegaard, despair is more than the absence of ground projects that convey meaning

34 For Schillerian criticism of Kantian ethics and its influence on Hegel and Kierkegaard, see Stern 2012, Ch. 4 and pp. 193 – 199. For Kierkegaard and German Romanticism, see Behler 1997; Bohrer 1987, pp. 62 ff.; Tjønneland 2004, esp. Ch. 2; Stewart 2003, pp. 170 – 181.

35 It should be clear that this argument only tries to make plausible that we should evaluate ourselves in moral terms; the argument does not try to show that doing so amounts to succeeding in fulfilling the ethical task (Davenport 2001b, p. 287). It is perfectly possible, of course, to recognize the authority of moral standards without living up to these standards.

to our lives. Kierkegaard actually claims that despair takes the form of being in conflict with oneself by having *two wills* that are inconsistent with one another. Kierkegaard writes, "everyone in despair has two wills, one that he futilely wants to follow entirely, and one that he futilely wants to get rid of entirely." (SKS 8, 144; UD, 30) Whereas the despair of possibility (futilely) wants possibility without necessity, the despair of necessity (futilely) wants necessity without possibility. The upshot is that it is only by willing the good unconditionally that one can will one thing, and therefore be in agreement with oneself and avoid despair (SKS 8, 139 f.; UD, 24; cf. Grøn 1997, pp. 261 f.). The real choice then stands between willing the good unconditionally and willing it to some degree only. Whereas the ethicist and the religious strive for the former, the aesthete can be said to settle for the latter.

But why does the aesthete despair or why can he be said to be in despair? The point seems to be that the aesthete has two different wills that cannot be reconciled. On the one hand, the aesthete is ruled by non-moral incentives and principles. We have seen Kierkegaard describing these in terms of sensuousness, self-love, self-interest, and happiness. The aesthete, then, is someone who is ruled by sensuousness, so that rationality and reflection serve sensuousness instead of morality (something that appears to make the aesthete heteronomous in the Kantian sense). On the other hand, the aesthete is not a mere natural being who could not have prioritized differently. The aesthete is not some animal that cannot be held responsible for his acts, since he has freely chosen to prioritize sensuousness over morality. However, subsuming morality under sensuousness means that morality is conditional on non-moral incentives or principles. This means that the aesthete acts morally in a very limited sense, that he, for instance, loves himself and his neighbor when he feels like it, but not all of the time.

However, this is deeply problematic since morality, by its very nature, requires unconditional and universal compliance. For if the will were to compromise on morality as the aesthete does, it would partially affirm its nature and partially affront it. It would partially express its essence and partially violate it, allowing itself to be determined sometimes by aesthetic standards and sometimes taking morality to be of absolute worth. But as Seiriol Morgan has pointed out, in trying to do so,

> the will would actually fail to achieve in any measure any of the things it half-heartedly attempted to commit to. You do not live up to the demands of morality at all by committing yourself to do so to a certain extent, and you cannot appreciate the dignity of humanity if

> you resolve to respect it only now and then. Rather, this would just show that you had failed to grasp the importance of any of these things in the first place.[36]

Throughout his writings, Kierkegaard argues that we cannot be entirely indifferent towards existence or life. As long as we are conscious, we have to adopt at least some kind of attitude towards our lives, selves, and surroundings. Even someone who suffers greatly will have to relate to this suffering by adopting some kind of attitude towards it. Put in Sartrean terms, this means that we are condemned to be free, since we have to exercise our freedom by choosing. Even the choice not to choose is a choice, Kierkegaard famously reminds us. This point may be stated more precisely by using Kantian terms (cf. Irwin 2011, vol. 3, p. 301): Our will (*Willkür*) has to incorporate incentives (*Triebfedern*) into our maxims (dispositions) and to posit ends. In order to follow inclinations, we must freely (spontaneously) incorporate these into our maxim instead of other incentives, since inclinations must be taken as an appropriate basis of action. And in order to pursue an end, we must have freely chosen it as our end.[37]

The next step then is to argue that we are responsible for our choices, at least insofar as we could have chosen otherwise or insofar as we could consider an inclination or an end to be appropriate to act on because it can be supported by reasons. This step makes it possible to blame someone for choosing incorrectly or for adopting the wrong kind of attitude. This point is usually seen as uncontroversial, and Kant and Kierkegaard both accept it.

The final step is to argue that my choice needs to be consistent with the choice of others so that my freedom does not undermine the freedom of others but rather promotes it. Basically, the reason for this is that I am dependent on others, since my self-consciousness, self-relation, and rationality are dependent on others. We therefore not only need to passively respect the freedom of others

---

36 Morgan 2005, pp. 96 f. This is the doctrine of moral rigorism associated with Kant. Unlike Rudd, I read Kierkegaard as a rigorist relying on a Kantian notion of ethics. The reasons for this are threefold: First, there is clear textual evidence for Kierkegaard's rigorism. Second, rigorism seems preferable to latitudinarianism, something that has been argued by Allison (1995, pp. 146 – 152), Firestone/Jacobs (2008, pp. 127 – 133) and Morgan (2005, pp. 96 f.). Finally, there are strong Kantian elements in Kierkegaard's ethics (Fremstedal 2014).

37 Allison comments: "I cannot conceive of myself as [...] [a rational] agent without regarding myself as pursuing ends that I frame for myself and that I regard as rational to pursue. Correlatively, I cannot conceive of myself as such an agent without assuming that I have a certain control over my inclinations, that I am capable of deciding which of them are to be acted upon (and how) and which resisted. These are, as it were, necessary presuppositions for all who regard their reason as practical. Kant indicates this in the *Groundwork* by suggesting that we cannot act except under the idea of freedom [...]" (Allison 1995, p. 41)

by refraining from violating it, but we also need to actively promote the freedom of others. Even though the principle of right may help with the former, it seems that moral obligations are needed to secure the latter. If these broadly Kantian points are accepted, then it seems to follow that we have to accept the necessity of moral restrictions insofar as we are free and conscious agents. The next section will go beyond these points by elaborating on our dependency on others.

Intersubjectivity, Love, and Emotions

Works of Love presents an interesting argument why the aesthete is in despair which supplements the arguments we have considered so far. *Works of Love* argues that hope without the moral duty to love one's neighbor is false, so that the real alternative to neighbor-love is despair. *Works of Love* proceeds by discussing a case where I hope for myself while giving up others by viewing them as hopeless (SKS 9, 253 – 256; WL, 254 – 256). However, hoping only for myself involves conceiving of hope and the good as something private that does not concern my relationship to others, as if I have a future of my own without others or as if what is good for me is entirely unconnected to what is good for others. Kierkegaard argues convincingly that by hoping in this way I fail to appreciate the extent to which I am dependent upon others. If there is no hope for others, then there cannot be any hope for me either, since I am dependent upon others. If they are trapped in hopelessness, this must also hold for me, even if I do not realize it myself. In this sense, I can be trapped in despair or hopelessness without realizing it.

Kierkegaard's point is that hoping for oneself must involve hoping for others, hoping for society (SKS 9, 253 f., 248; WL, 253 f., 248). Kierkegaard stresses that neighbor-love takes upon itself the work of hope, the task of hoping for others: "love is [...] the middle term: without love, no hope for oneself; with love, hope for all others—and to the same degree one hopes for oneself, to the same degree one hopes for others, since to the same degree one is loving." (SKS 9, 259; WL, 260) Love thus connects hope for oneself with hope for others, transforming the object of hope into something universal, arguably an ethical commonwealth or invisible church.[38] The upshot is not only that there must be some connection

38 One may object that one is not dependent on all human beings but only on some. However, whom I depend upon in different contexts seems contingent. There does not seem to be a principled reason that prevents me from being or becoming dependent on anyone in particular. Still, Kierkegaard's point is not mainly that I may find myself being dependent upon a stranger or an enemy. Put in Apelian or Habermasian terms, Kierkegaard is concerned rather with how actual

between what is good for me and what is good for others but also that one is trapped in despair without the moral obligation to love one's neighbor.

Kierkegaard appears to make use of the broadly Hegelian idea that self-consciousness presupposes intersubjectivity, and that the self-relation is mediated by the other (cf. Furtak 2005, pp. 74, 99; Grøn 1997, Ch. 5). Although Kierkegaard does not fully accept Hegel's ethics of recognition, he does agree that I am dependent on others and that this dependency implies moral restrictions. I must behave so that my freedom does not undermine the freedom of others but rather promotes it by letting the other stand on his own as an individual different from me. The central claim here is that we need something to mediate between the self and the other, and that this mediating principle needs to be ethical. Whereas Hegel describes this principle in terms of moral recognition, Kierkegaard describes it in terms of neighbor-love.[39]

---

discourses (performatively) presuppose an ideal discourse. Put in ecclesiastical terms, he is concerned with how the individual depends on the invisible (true) church rather than the visible church. The Kantian parallel to this seems to be Kant's claim, in Book III of *Religion within the Boundaries of Mere Reason*, that individual struggle against moral evil requires an ethical commonwealth that makes the victory of the good principle over the evil one possible. Hare elaborates on this by saying that "His argument is that we will have ends which require the help of others if we are going to reach them [...]. We are linked together by our needs and abilities into a single unit, or kingdom, which we must be prepared to will into existence as a whole. It contains our needs (for even in the true church we will be creatures of need), and it contains other people with the developed abilities to meet our needs; but it also contains the needs of others, and our developed abilities to meet their needs." Hare 2002b, p. 265.

39 A different but related argument is sketched in *Fear and Trembling*. *Fear and Trembling* argues that I am dependent on the understanding of others for a right understanding of what to do. More specifically, it is argued that I am likely to overlook relevant arguments unless I communicate openly with others (SKS 4,177; FT, 87). Based on this, Vittorio Hösle (1992, p. 7) has argued that Kierkegaard anticipates the central idea of Habermas's discourse ethics. Put in Kantian terms, the touchstone for assessing the objectivity of subjective judgments (claims about something being true) is whether others agree. Kant says that we cannot do without the understanding of others because such an understanding is an *external criterion* for truth. Without this criterion, we could not test the correctness of our own judgments, and hence we would be at the mercy of mistakes. This criterion means that the rational validity of judgments depends on it being possible to *communicate or share* them universally (Fremstedal 2009, pp. 27 f.). Put in contemporary terms, the point seems to be that we are dependent on a practical discourse for understanding what we ought to do. And this practical discourse, like intersubjectivity more generally, presupposes certain ethical norms. Kierkegaard's ethicist formulates this by saying that personal and civic virtues are interdependent, so that I cannot have any personal virtues without also having civic or social virtues (SKS 3, 249; EO2, 262). The ethicist distinguishes between personal virtues (courage, valor, temperance, and moderation) that are necessary for self-development and civic virtues (notably justice) that are necessary for participation in social life.

Both the aesthete and the ethicist are deeply concerned with love as an emotion (or a passion). The aesthete focuses on romantic love, whereas the ethicist focuses on marriage as the paradigm case for the ethical stage.[40] The aesthete believes in love as an experience that makes life beautiful and interesting, seeing marriage and its duties as incompatible with the freedom and spontaneity required by genuine love (Davenport 2001b, pp. 91f.). Kierkegaard suggests, however, that there is some continuity between the different forms of love. *Stages on Life's Way*, for instance, describes different forms of love in a manner reminiscent of Plato's famous description of different forms of love in the *Symposium* (cf. Furtak 2005, pp. 103f.). Whereas the aesthete sees love as a mere feeling, the ethicist and Kierkegaard seem to approach it as a virtue with affective and emotional aspects that can, and ought to, be cultivated.

The ethicist argues that even though we have first-order desires and states, these can only have (lasting) significance by being actively endorsed and guided by practical rationality involving intersubjective standards of assessment. More specifically, first-order desires only acquire real importance if we are ethically committed by relating to what happens to us, either by identifying with or distancing ourselves from first-order desires. This can be done either by viewing first-order desires as appropriate or as inappropriate, as something we ought or ought not to act on, based on the merits of different options, or on reasons that hold irrespective of the strength of our inclinations (Irwin 2011, vol. 3, pp. 299f.). We thus need to introduce the idea of a rational choice that is based on the merits of different options, not just on inclinations and desires. This means that we enter the area of good and evil as features to be considered in a free choice (Irwin 2011, vol. 3, p. 300).

The aesthete experiences different emotions and desires, but he does not give his assent to them by actively embracing their significance or by endorsing them. He lacks second-order desires that make it possible to develop and cultivate first-order desires and to develop a long-standing attitude. He hardly views himself as the owner of his inclinations and desires, and he does not take re-

---

Even if there is some uncertainty how far Kierkegaard actually developed this point, it nevertheless seems to be a valid point that could have developed by him. Still, it seems that some of this point is perhaps better, or more systematically, developed by Kant, Hegel, and Habermas than by Kierkegaard, even though Kierkegaard goes beyond these thinkers by analyzing despair.

40 The ethicist makes use of the Fichtean idea that marriage is a step on the path to becoming an ethically developed person and that the love relation is nature's way of overcoming itself and pushing us toward becoming ethical beings. Marriage is therefore considered a duty (something both Kant and Kierkegaard were opposed to). See Kosch 2006b, pp. 270 – 273. See also Hannay 2006, pp. 94 – 101.

sponsibility for them, although they do make up the basis of his decisions, ends, and actions. He is free in the sense of being independent of his inclinations, but he does not recognize or affirm his freedom like the ethicist does (Irwin 2011, vol. 3, pp. 301, 304). The reflective aesthete takes up the perspective of a spectator towards his own emotions and his own life. By doing this, he denies that he is already involved in life and therefore responsible.[41]

The aesthete is detached from his ends, not because he chooses them without energy or dedication, but because he sees them as external to himself, insofar as they are objects of inclination that are purely accidental to him (Irwin 2011, vol. 3, p. 299). The aesthete would be less detached from his ends if he could regard them as appropriate for him, as the sort of agent he is, because they represent the type of ends that he ought to choose irrespective of the strength of his inclinations (Irwin 2011, vol. 3, p. 299). Terence Irwin writes:

> Since we regard ourselves as continuing selves; and think it right, irrespective of the strength of our desires, to plan for our continuing selves, we can also see—though we may not see—that a purely aesthetic attitude to ourselves cannot satisfy us. If we treat our ends as matters of mere inclination, we do not ask the questions that, as continuing agents, we recognize as legitimate, about whether we have reason to pursue this end rather than another. The aesthetic outlook does not fit the self that adopts it. (Irwin 2011, vol. 3, p. 299)

Irwin concludes that the aesthetic agent is liable to despair because aesthetic agency presupposes some basis of non-aesthetic agency (Irwin 2011, vol. 3, pp. 299 f.). The aesthete thinks of himself as a particular continuing self that is free but this self does not fit aesthetic agency, which is ruled by inclinations and desires that are accidental and external.

The ethicist argues that romantic love needs to be both endorsed and restricted in marriage. The idea is that romantic love is transfigured in marriage so that love's needs are completed and fulfilled. On this view, marriage is not an alien imposition on romantic love, but something that makes it possible for romantic love to develop and endure. The ethicist proceeds by arguing that love itself wants to be strengthened, since it wants to ensure that love will last. Even in the absence of a marriage ceremony lovers therefore swear faithfulness to each other in the name of something perceived to be higher (e. g. moon, stars, father's ashes) so as to bind themselves (SKS 3, 61 f.; EO2, 56; Davenport 2001b, pp. 91–94). This indicates that love itself seeks moral commitment

41 Furtak 2005, pp. 58 f., 62 f., 66, 79. For the interpretation of the reflective aesthete as a sophisticated wanton, see Rudd 2012, pp. 80–85; Davenport 2001b, pp. 294 f.

(SKS 3, 61, 66, 144–147; EO2, 56, 60f., 146–149).[42] The ethicist presents a somewhat Hegelian (and Schillerian) argument to the effect that moral obligation is not opposed to love, as has been pointed out by existing scholarship (cf. Stern 2012, pp. 190–199). At this point the ethicist stresses that moral duty should not be interpreted as something external that is opposed to my inner being, but rather seen as something that expresses my true being (SKS 3, 242f.; EO2, 254f.). Freedom is therefore seen as realized *in* moral and social commitment.

The more general point, however, is that there is no free lunch. Things cannot have (lasting) importance or meaning if they do not imply some commitment or obligations. Emotions that are not actively endorsed and regulated are merely episodic sensations without meaning. The aesthete lives in a world of fleeting and abbreviated emotions, lacking emotional integrity (Furtak 2005, pp. 59, 65). He may consider emotions and passion to be the deepest part of the human being; but these are wild and unruly as long as the aesthete does not have any definitive aim or end (as the ethicist does).[43] Without an active endorsement of emotions, these will disintegrate into mere fragments and the aesthete will be ironic and indifferent towards his own life. Furtak elaborates:

> He avoids taking anything seriously, and thereby guards himself against the emotional risk of being more than ironically involved. And the fragmentary nature of his temporal existence also keeps him from occupying any role that requires sustained care: he can be a dilettante but not a devoted artist, a temporary acquaintance but not a loyal friend. [...] Rather than letting his episodic emotions grow into longstanding attitudes, the aesthete lets them weightlessly pass away, so that both joy and torment end up meaning nothing. (Furtak 2005, pp. 68, 79)

Conclusion

The points sketched above are extracted from Kierkegaard's writings, notably *Either/Or*.

Rather than proving why we should be moral, Kierkegaard can be said to provide strong motivation for a transition from the aesthetic to the ethical

42 Kierkegaard himself has a somewhat less optimistic take on this in *Works of Love*. He argues that love's need for obligation show that love is dimly aware that it is insufficient by itself; love is insecure, anxious about the possibility of change, that love may vanish or change. As a result, love needs moral obligation (SKS 9, 40, 73; WL, 32f., 66). Kierkegaard's claim that difficulties remain with B's notion of marriage (SKS 7, 167; CUP1, 181) need not undermine B's general attempt to criticize the aesthete or his attempt to explain why we should be moral.

43 Furtak 2005, p. 77 referencing Johannes Sløk.

stage by arguing that despair can only be overcome if we choose the ethical. *Either/Or* portrays the ethical as inescapable, so that we can speak of the intrusion of the ethical (cf. Grøn 1997, pp. 261f.; Evans 2009, pp. 87–89). MacIntyre is therefore mistaken in claiming that Kierkegaard's existential choice between the aesthetic and the ethical is criterionless like the radical choice of Sartre (Davenport 2001a).

Perhaps the most characteristic feature of Kierkegaard's work, as compared to his predecessors, is the central role despair plays in it. Kierkegaard went beyond his predecessors by analyzing the importance of despair and hope for moral agency, offering a systematic analysis of despair that makes extensive use of (moral) psychology and phenomenology.[44] Still, it might seem that Kierkegaard's general methodology is perhaps stronger than the specific arguments. The arguments are typically incomplete and sketchy, standing in need of interpretation and reconstruction.[45]

The renewed interest in Kierkegaard over the last decades has led both to new historical research and attempts to use Kierkegaard in contemporary debates. The present paper belongs mainly to the latter category and focuses on reasons for seeing the human task as the ethical task, seeing the choice of oneself as the choice of the ethical. Even if the arguments considered are somewhat incomplete or equivocal, they still seem to make plausible that we need to evaluate ourselves in moral terms. It seems fair therefore to conclude that Kierkegaard comes at least some way towards answering the question "Why be moral?" – although his approach is more relevant to deontological ethics and virtue ethics than consequentialism.

Kierkegaard's work is so rich and multi-faceted that it has the potential for adding something valuable to contemporary discussions, as is exemplified by the work of Rudd, Davenport, Furtak, and others. What makes Kierkegaard's work interesting are not only its arguments and dialectics but also its vivid literary descriptions and examples as well as its use of phenomenology and psychology. I agree with Davenport that, in the contemporary context, Kierkegaard may be seen as allied with MacIntyre, Korsgaard, and Taylor against Williams in thinking that non-moral caring involves implicit rational commitment to ethical values, whether or not we recognize it or like it. The connection that Kierkegaard draws between earnest purpose in life and choosing the ethical is controversial

44 Cf. Stokes 2010, pp. 7f.; Hannay 2006, p. 4. Regarding phenomenology, see Grøn 1997, pp. 137–142.

45 Theunissen (1993, pp. 13, 108) has pointed out that there exist relatively few attempts to defend or reconstruct Kierkegaard's theory in a rational or argumentative manner.

and provocative, but it should nevertheless be taken as seriously as similar views voiced by contemporary thinkers (cf. Davenport 2012, pp. 130, 122).

Bibliography

Allison, Henry (1995): *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
Annas, Julia (1993): *The Morality of Happiness*. Oxford: Oxford University Press.
Behler, Ernst (1997): "Kierkegaard's *The Concept of Irony* with Constant Reference to Romanticism". In: *Kierkegaard Revisited*, Niels Jørgen Cappelørn and Jon Stewart (Eds.), Berlin: de Gruyter, pp. 13 – 33.
Bohrer, Karl Heinz (1989): *Die Kritik der Romantik. Der Verdacht der Philosophie gegen die literarische Moderne*. Frankfurt am Main: Suhrkamp.
Davenport, John (2001a): "The Meaning of Kierkegaard's Choice between the Aesthetic and the Ethical: A Response to MacIntyre". In: *Kierkegaard after MacIntyre*, John Davenport and Anthony Rudd (Eds.), Chicago: Open Court, pp. 75 – 112.
Davenport, John (2001b): "Towards an Existential Virtue Ethics: Kierkegaard and MacIntyre". In: *Kierkegaard after MacIntyre*, John Davenport and Anthony Rudd (Eds.), Chicago: Open Court, pp. 265 – 323.
Davenport, John (2012): *Narrative Identity, Autonomy, and Mortality: From Frankfurt and MacIntyre to Kierkegaard*. London: Routledge.
Evans, C. Stephen (2006): *Kierkegaard's Ethics of Love: Divine Commands and Moral Obligations*. Oxford: Oxford University Press.
Evans, C. Stephen (2009): *Kierkegaard: An Introduction*. Cambridge: Cambridge University Press.
Ferreira, M. Jamie (2009): *Kierkegaard*. Oxford: Wiley-Blackwell.
Firestone, Chris/Jacobs, Nathan (2008): *In Defense of Kant's* Religion. Indianapolis: Indiana University Press.
Fremstedal, Roe (2009): "Critical Remarks on 'Religion in the Public Sphere' – Habermas between Kant and Kierkegaard". In: *Etikk i praksis. Nordic Journal of Applied Ethics* 3, pp. 27 – 47.
Fremstedal, Roe (2014): *Kierkegaard and Kant on Radical Evil and the Highest Good: Virtue, Happiness, and the Kingdom of God*. Basingstoke: Palgrave Macmillan.
Furtak, Rick Anthony (2005): *Wisdom in Love: Kierkegaard and the Ancient Quest for Emotional Integrity*. Notre Dame: University of Notre Dame Press.
Grøn, Arne (1997): *Subjektivitet og negativitet: Kierkegaard*. Copenhagen: Gyldendal.
Hannay, Alastair (2006): *Kierkegaard and Philosophy: Selected Essays*. London: Routledge.
Hare, John (2001): *God's Call: Moral Realism, God's Commands, and Human Autonomy*. Grand Rapids: Eerdmans.
Hare, John (2002a): *Why Bother Being Good?* Downers Grove, Ill.: InterVarsity Press.
Hare, John (2002b): *The Moral Gap: Kantian Ethics, Human Limits, and God's Assistance*. Oxford: Clarendon Press.
Hatfield, Gary (2005): "The History of Philosophy as Philosophy". In: *Analytic Philosophy and History of Philosophy*, Tom Sorell and G. A. John Rogers (Eds.), Oxford: Clarendon Press, pp. 83 – 128.

Horn, Christoph (1998): *Antike Lebenskunst. Glück und Moral von Sokrates bis zu den Neuplatonikern*. Munich: C. H. Beck.
Hösle, Vittorio (1992): "KAN ABRAHAM REDDES? OG: KAN SØREN KIERKEGAARD REDDES? Et hegelsk oppgjør med 'Frygt og Bæven'". In: *Norsk filosofisk tidsskrift* 27, pp. 1–26.
Irwin, Terence (2011): *The Development of Ethics: A Historical and Critical Study*, vols. 1–3, Oxford: Oxford University Press.
Kekes, John (2013): "Meaning and Narratives". In: *On Meaning in Life*, Beatrix Himmelmann (Ed.), Berlin: de Gruyter, pp. 65–82.
Knappe, Ulrich (2004): *Theory and Practice in Kant and Kierkegaard*. Berlin: de Gruyter.
Kosch, Michelle (2006a): *Freedom and Reason in Kant, Schelling, and Kierkegaard*. Oxford: Oxford University Press.
Kosch, Michelle (2006b): "Kierkegaard's Ethicist: Fichte's Role in Kierkegaard's Construction of the Ethical Standpoint". In: *Archiv für Geschichte der Philosophie* 88, pp. 261–295.
Lübcke, Poul (1991): "An Analytic Interpretation of Kierkegaard as a Moral Philosopher". In: *Kierkegaardiana* 15, pp. 93–103.
Morgan, Seiriol (2005): "The Missing Formal Proof of Humanity's Radical Evil in Kant's *Religion*". In: *The Philosophical Review* 114, pp. 63–114.
Nadler, Steven (2005): "Hope, Fear, and the Politics of Immortality," in *Analytic Philosophy and History of Philosophy*, Tom Sorell and G. A. John Rogers (Eds.), Oxford: Clarendon Press, pp. 201–217.
Poole, Roger (1997): "'My wish, my prayer': Keeping the Pseudonyms Apart. Preliminary Considerations". In: *Kierkegaard Revisited: Proceedings from the Conference "Kierkegaard and the Meaning of Meaning it," Copenhagen, May 5–9, 1996*, Niels Jørgen Cappelørn and Jon Stewart (Eds.), Berlin: de Gruyter, pp. 156–176.
Roberts, Robert C. (1998): "Existence, Emotion, and Virtue: Classical Themes in Kierkegaard". In: *The Cambridge Companion to Kierkegaard*, Alastair Hannay and Gordon D. Marino (Eds.), Cambridge: Cambridge University Press, pp. 177–206.
Rudd, Anthony (2001): "Reason in Ethics: MacIntyre and Kierkegaard". In: *Kierkegaard after MacIntyre*, John Davenport and Anthony Rudd (Eds.), Chicago: Open Court, pp. 131–150.
Rudd, Anthony (2005): *Kierkegaard and the Limits of the Ethical*. Oxford: Clarendon Press.
Rudd, Anthony (2012): *Self, Value, and Narrative: A Kierkegaardian Approach*. Oxford: Oxford University Press.
Stewart, Jon (2003): *Kierkegaard's Relations to Hegel Reconsidered*. Cambridge: Cambridge University Press.
Stokes, Patrick (2010): *Kierkegaard's Mirrors: Interest, Self, and Moral Vision*. Basingstoke: Palgrave Macmillan.
Theunissen, Michael (1991): *Das Selbst auf dem Grund der Verzweiflung. Kierkegaards Negativistische Methode*. Frankfurt am Main: Anton Hain.
Theunissen, Michael (1993): *Der Begriff Verzweiflung. Korrekturen an Kierkegaard*. Frankfurt am Main: Suhrkamp.
Thuborg, Anders (1951): *Den Kantiske periode i dansk filosofi, 1790–1800*. Copenhagen: Gyldendal.
Tjønneland, Eivind (2004): *Ironie als Symptom. Eine kritische Auseinandersetzung mit Søren Kierkegaards* Über den Begriff der Ironie. Bern: Peter Lang.

Richard Eldridge

Acknowledging the Moral Law

> Knowledge is in the end based on acknowledgment.
> Ludwig Wittgenstein, *On Certainty*, §378

I

Once upon a time, almost forty years ago, when it was still common for graduate students to take qualifying examinations in various subfields of philosophy, there was—so the lore among the graduate students had it—a standard, expected reply to the question "Why be moral?" One was supposed to divide one's answer into two cases. If the question were understood as asking "What self-interested, prudential reasons are there to be moral?", then the answer is "obviously none; often enough one will be either inconvenienced or otherwise disadvantaged by doing what one has, for example, promised to do." Alternatively, if the question were understood as asking "What *moral* reasons are there to be moral?", then the answer is "obviously whatever moral reasons one has; it is essential to the distinctive force of moral reasons that they cannot be reduced to other sorts of reasons."[1]

In retrospect, both these lines of response to the question "why be moral?" are, at best, evasive, and we have, happily, to some extent learned to think more deeply about how moral reasons make claims on us. In large part, this new thinking has been spurred by Bernard Williams's questioning of the absolute authority of distinctively moral reasons (to do one's duty, to keep one's promises, etc.) and his urging of the importance in contrast of the *ethical:* a set of open-ended, plural, nonsystematizable considerations about values—moral, prudential, political, intellectual, aesthetic, and otherwise—that one might articulate and to which one might appeal more or less reasonably and improvisatorially in order to make sense of one's life.[2] As Derek Parfit puts it in endorsing Wil-

1 Notably, Bernard Williams characterizes Kantian morality in general as making use of this dilemma and as then, vacuously as he sees it, embracing its second horn. See Williams 1985, pp. 54–55.

2 Alasdair MacIntyre's questioning in *After Virtue* of the authority of modern liberal morality and his urging of a modified form of virtue ethics have also been influential (MacIntyre 1981).

liam's recommendation that we take up the Socratic question, "how should one live?,"[3]

> We should ask what we have reasons to care about, and to try to achieve. [...] Reasons are, I believe, fundamental. Something matters only if we have reason to care about this thing. It would have great importance if morality did not in this sense matter, because we had no reason to care whether our acts were right or wrong. To defend and explain morality's importance, we can claim and try to show that we do have such reasons. Morality might have supreme importance in the reason-implying sense, since we might always have decisive reasons to do our duty, and to avoid acting wrongly. But if we defend morality's importance in this way, we must admit that the deepest question is not what we ought morally to do, but what we have sufficient or decisive reasons to do.[4]

While Parfit here displays an admirable forthrightness in taking seriously Williams's question about the authority of morality—in asking for *reasons* for being moral—his own response to this question arguably suffers from both insensitivity to available possibilities of practical reasoning and incoherence.[5] Worse yet, Parfit, in his relentless hunt for binding reasons to be moral, skates over the most important insights of Williams's philosophical anthropology that make the question about the authority of morality significant for us, especially Williams's sense that we are fragile beings for whom achievements of the good, even sometimes life itself, are hostage to fortune and dependent on our relations with specific others. As Williams observes, we are now in some respect closer to the Greeks—and perhaps even closer to them now than we have been for the last forty years or two millennia—than has often been supposed, particularly in feeling the legitimate force of the claims of personal relations and in feeling bound up in processes we can neither escape nor master, all while lacking the comforts of belief in literal bodily resurrection after death.[6]

---

3 Williams 1985, p. 4. Williams goes on to remark "I shall not try to define what counts exactly as an ethical consideration; ... it does no harm that the notion is vague" (p. 7). Roughly, ethical considerations turn out to be for Williams *any* considerations—moral, economic, political, aesthetic, prudential, sexual, familial, and so on—that bear seriously on how one should live *as* the person who one most deeply is.

4 Parfit 2011, II, p. 444, I. p. 148.

5 In reviewing Parfit's book, Samuel Freeman notes i) that Parfit regards the only options in metaethics as either extreme rationalism or extreme subjectivism, thus missing a range of available intermediate stances, and ii) that Parfit combines a Kantian concern for absolute human rights with utilitarian concern for human welfare in a way that is never made fully clear in relation to hard cases where these concerns might clash. (Freeman, 2012, pp. 52 – 54.)

6 See Williams 1993, p. 166.

This philosophical anthropology has important consequences for how we should think about philosophy as a discipline or activity. We should abandon what Williams called the bogus "aspiration to a total critique,"[7] give up the attempt to arrive at a view from nowhere, as if we might, by doing so, make ourselves immune to fortune and certain of our own moral worth and purity. Instead, we should accept our embeddedness in natural processes and personal and social relations. Philosophy is, or should be if it is honest, as Williams puts it, a matter of trying to "make the best sense of our life, and so of our intellectual activities, *in the situation in which we find ourselves*."[8] Normative relations, including for example semantic relations, are instituted by us in the course of our complex practical lives, and they are subject to historical variation. (Williams cites with approval R. G. Collingwood's remark that certain practitioners of putatively ahistorical and absolute analysis of concepts proceed by, as it were, "translating the Greek word for a trireme as a 'steamship' and then complain[ing] that the Greeks had a defective conception of a steamship."[9]) Philosophy should not be disjoined from history; making sense of semantic and other normative phenomena must involve paying reflective attention to concept-words in actual use and subject to change. We should not suppose that our own stocks of concepts could or should be perfectly in order.

Pointedly, we should not suppose that we are morally superior to the Greeks. Among them, "slavery, in most people's eyes, was not just, but necessary."[10] Most Greeks could not imagine a world without it. Rightly, we take this to be a failing on their part. But our own attitudes toward economic immiseration and degradation are not so dramatically different, insofar as many of us unreflectively accept radical inequalities that we regard as in principle unjust as the necessary price of significantly rising GDP. Worse yet, under the banner of unrestricted property rights, freedom of contract, and economic efficiencies, some of us celebrate radically immiserating and degrading property arrangements as just, even though all property holdings are matters of institutional arrangement, not of natural right and clean historical descent. In some ways, the Greeks confronted the difficult facts of their historical situation more forthrightly than we sometimes confront ours.

Nor is our moral psychology obviously in much better shape than that of the Greeks, since there is no possibility of pure guilt or the verdict of conscience simply as such, arising from an inner confrontation with the bare moral law. Rather,

7 Williams 1993, p. 159.
8 Williams 2008, p. 182; emphasis added.
9 Williams 2008, p. 181.
10 Williams 1993, p. 117.

as Williams argues, drawing on Nietzsche, guilt before a sense of what is morally required of one is itself a refined and abstracted form of shame. As Williams puts it, "shame can understand guilt, but guilt cannot understand itself. ...Only shame [not guilt] can ... help one to understand one's relations to [wrongs one has done or might do], because it [unlike guilt] embodies conceptions of what one is and of how one is related to others."[11] And this is because, as Nietzsche together with Freud (on the formation of the superego) have taught us, the voice of conscience that proclaims guilt before the moral law is itself the voice of an internalized other that has been modeled on and abstracted from some actually existing other, with authoritative standing, within social-ethical life.[12] Any other picture of how we come to have senses of duty and obligation—any picture that posits a pure practical reason within, or responsiveness to the voice of God, capable of operating quite independently of the internalization of admiration, respect, fear, and so on toward actually existing others—is simply horribly unrealistic. True, we can abstract and generalize away from the evaluative stances of the particular others we have encountered, partly because we encounter many others with many distinct evaluative stances. But however far it goes, such abstraction and generalization continue to bear some traces of some actually existing encountered and internalized others. In Williams's summary formulation,

> By giving through the emotions a sense of who one is and of what one hopes to be, [shame, along with its later abstracted and derivative form, guilt] mediates between act, character, and consequence, and also between ethical demands and the rest of life. Whatever it is working on, it requires an internalized other, who is not designated merely as a representative of an independently identified social group, and whose reactions the agent can respect. After some time, this figure does not merely shrink into a hanger for those same values but embodies intimations of a genuine social reality—in particular of how it will be for one's life with others if one acts in one way rather than another.[13]

Finally, Williams is right that consequences of actions, even unintended ones, sometimes matter. It is at least sometimes appropriate to feel regret for what one has done, even if one did not intend to do it or even intended specifically not to do it. As Hegel trenchantly puts it, "'The stone belongs to the devil when it leaves the hand that threw it.' When I act, I allow for bad luck, so it has a right over me and is an existence of my own willing."[14] We would do better to learn to live with this fact rather than denying it. Or as Williams puts it, "we

11 Williams 1993, pp. 93 – 94.
12 Williams 1993, p. 84.
13 Williams 1993, p. 102.
14 Hegel 1991, p. 148.

know that in the story of one's life there is an authority exercised by what one has done, and not merely by what one has intentionally done."[15]

Given all this—the fragility of goodness, the properly historical character of philosophical thinking about normativity, the developmental priority of shame over guilt, and the fact of moral luck—why should we not, as Williams urges, content ourselves with a looser "conception of the ethical that understandably relates to us and our actions the demands, needs, claims, desires, and, generally, the lives of other people"[16] in various ways? Why, if at all, should we embrace what Williams calls "the purity of morality" with its "'must' that is unconditional and *goes all the way down*"?[17]

II

Throughout both *Ethics and the Limits of Philosophy* and *Shame and Necessity*, Kantianism, with its sterile and impotent conception (as Williams sees it) of distinctively moral reasons that absolutely bind us from nowhere—in contrast with broader ethical reasons that arise within the course of life—is the official target.[18] At first blush, moreover, Kant's account of the authority of morality can indeed seem both inhuman in being insensitive to the conduct of practical life on the ground and distressingly dogmatic. Not only does he simply claim that "the moral law is given, as it were, as a fact of pure reason of which we are a priori conscious and which is apodictically certain,"[19] he also describes awareness of this fact as a matter of "conscientiousness ... as accountability to a holy being (morally lawgiving reason) distinct from us yet present in our inmost being."[20] Hence he is at least committed to a doctrine of a split in the human subject between an empirical part and a pure rational part. At best, this split can seem obscure. Hence to make Kantian distinctively moral reasons *count* for us would seem to have to mean something like training our defective, empirical, desiring,

15 Williams 1993, p. 69.
16 Williams 1985, p. 12.
17 Williams1985, pp. 195, 188.
18 See the pages from (Williams 1985) cited in note 1 and also p. 104: "Hegel admirably criticized the 'abstract' Kantian morality and contrasted it with the notion of *Sittlichkeit*, a concretely determined ethical existence that was expressed in the local folkways, a form of life that made particular sense to the people living in it." As Williams immediately adds, his own position is a Hegelian defense of *Sittlichkeit* stripped, however, of any reference to teleology.
19 Kant 1996b, p. 177.
20 Kant 1996c, p. 561.

relationship-embedded, and project-having nature *to be more at home with reasons that are indifferent to that nature*.[21] Rightly, that can seem a very tall order, perhaps too tall.[22]

It is, after all, true that human beings in the course of their individual developmental lives from infancy onwards find themselves having projects and desires and are situated within relationships that have force for them. Why then, if at all, should distinctively moral reasons, involving universal and impersonal respect for persons, function as a kind of standing counterforce that checks and corrects the commitments we already have?

If there is an answer to this question, it must involve seeing the emergence of distinctively moral commitments out of an enormous variety of developmental backgrounds, both individual and social-historical, as displaying a kind of path independence. That is, just as pebbles dropped from various positions above the top of a well may follow a variety of distinct paths, including bouncing off the sides of the well in various ways, they will nonetheless each end up at the bottom of the well.[23] This is, of course, due to gravity. Could—and should—the moral law have for us anything like the dispositive force of gravity, in bringing each of us to acknowledge its authority, despite our strikingly distinct individual and social paths? And here it is obvious that, while in some possible world, perhaps, this *could* happen, it scarcely seems likely as a matter of empirical fact that it has happened or is about to happen.

---

21 I owe this formulation to Robert Guay, who provided a series of well-considered, close, critical comments on an earlier draft of this essay and whose own defense of a Nietzschean ethical contextualism, along lines similar to those of Williams, was much on my mind in revising my argument.

22 Terry Pinkard, for example, describes "the notion of the 'fact of reason'" as "a restatement of the quasi-paradoxical formulation of the authority of the moral law itself, which seems to require a 'lawless' agent to give laws to himself on the basis of laws that from one point of view seem to be prior to the legislation and from another point of view seem to be derivative from the legislation itself," and he argues that this paradox can be overcome *only* by replacing Kant's appeal to pure practical reason in the individual with an account of the logic of doubling (*Verdopplung*) of agency among multiple subjects, each of whom is struggling to impose demands on others and to secure their recognition as reasonable (Pinkard 2002, pp. 59, 227). Karl Ameriks observes similarly that "if the source [of the moral law] is elevated into something even partly outside the world, then even when it is not personified as a literally separate and self-subsisting entity, it can still seem too odd and remote to explain why human agents should feel bound by it." (Ameriks 2012, pp. 150–51); see also p. 54, where Ameriks observes that "this can all sound too remarkable to be true."

23 I owe this useful analogy to Sam Baron, who suggested it during discussion of a version of this paper, subsequent to presenting it to the philosophy department of the University of Western Australia.

It is noteworthy that Kant himself held that consciousness of the moral law is not explicitly present in the consciousness of any human being from the moment of conception or birth. As he remarks in the *Foundations*, "innocence is indeed a glorious thing [though] it is sad that it cannot well maintain itself, being easily led astray."[24] Here innocence must consist in a time of life *before* one has become conscious of the moral law as binding normatively against the force of inclination. Likewise, in "Conjectural Beginning of Human History," Kant remarks, drawing on both *Genesis* and Milton,[25] that in the history of mankind "instinct ... must alone have guided the novice" well before "reason began to stir."[26] What, then, might the process of *coming to awareness* of the moral law *look like?* And what, if anything, might move us to think that this process is something that *should* take place within any individual's course of development?

We should not suppose, however, that a description of any such process will provide an empirical *explanation* of the origin, content, and authority of the moral law. It must, at least within a Kantian framework, be the case that the human subject autonomously exercises rational powers in coming to acknowledge the moral law. Nor will such a description provide an independent *justification* of the content and authority of the moral law, apart from a free act of rational acknowledgment that must be carried out by each agent. In this sense, the content and authority of the moral law cannot be proven to a moral skeptic.[27] The best one can do is to supply neither an empirical explanation nor an independent justification but rather, as Ameriks puts it, "just many layers of illuminating description" that capture *how* sometimes one may find oneself "in the situation of being able to maintain one's rationality only by [actively] acknowledging the pure practical law as compelling."[28]

In his "History" essay, Kant himself develops in some detail the idea of the moral law coming to have authoritative normative force for us in a path-independent way. At the end of the first paragraph of that essay, Kant observes that

> While [among individual human beings and even whole nations] each pursues its own aim in its own way and one often contrary to another, they are proceeding unnoticed, as by a

24 Kant 1959, p. 21.

25 For an extended account of Kant's adaptation of Milton's account of a procedure of succession (*Nachfolgung, Nachmachung*) as essential for coming to self-conscious maturity, see (Budick 2010).

26 Kant 2007, p. 165.

27 Hence there is something right about embracing the second horn of the dilemma with which this essay began.

28 Ameriks 2012, p. 181; compare the account of the acknowledgment of the moral law in Eldridge 1989, pp. 185 – 88.

> guiding thread [an einem Leitfaden], according to an aim of nature [Naturabsicht], which is unknown to them, and are laboring at its promotion.[29]

Here the guiding thread is explicitly the normative force with which human beings in various circumstances are drawn toward life on the ground according to the requirements of the moral law, no matter what diverse paths they may take toward this end and what other commitments they may also have. We are, as Kant puts it repeatedly in the *Religion*, each to live up to the "idea of a human being morally well-pleasing to God" ["Idee eines Gott moralisch wohlgefälligen Menschen"],[30] and we are to do so by creating first a lawful civic order in which each possesses the maximum degree of liberty compatible with the like liberty of all and second a kingdom of ends, that is, a moral culture of mutual respect and achieved concrete life according to reason.

Following this guiding thread—that is, moving toward life according to the moral law whose normative authority one accepts—is said, further, to take place "according to an aim of nature." In this context, "nature" *cannot* mean "the realm of empirical objects" or "nature as the object of study of the natural sciences," for nature in those senses has no aims or purposes [Absichten].[31] Instead Kant must mean something more like Spinozist *natura naturans* or "nature as God has made it to be purposive" (even where such purposes are not empirically discernible), in contrast with *natura naturata*. Nonetheless, one may still wonder: why *should* we be drawn normatively by the thought that we are participants in *natura naturans* or members of an intelligible world? The empirical world of ordinary objects is real enough and unavoidable, and we lack, both by Kantian lights and in fact, theoretical *knowledge* of any such higher order or world undergirding the ordinary world. Hence talk of a normatively dispositive noumenal nature of which we are members may seem to be little advance over sheer dogmatism, and such talk is, again, at best redescriptive of what we are implicitly and immanently committed to in virtue of our deliberative powers; it is neither an empirical explanation of morality nor an independent justification of it.

---

29 Kant 2007c, p. 108.

30 Kant 1996e, p. 105 and infra. Note, however, that to be well-pleasing to God just *is* to act from respect for the moral law; Kant, of course, does not put forward a divine command morality, and belief in (a thin, non-personal) God is derived from the requirements of the moral life rather than vice versa.

31 See Ameriks 2012, p. 271 for a useful account of ambiguities attaching to Kant's uses of "nature."

Is this redescriptive sketch of our powers and possibilities of acknowledging our membership in a noumenal world and accepting the authority of Kantian morality, from within divergent local circumstances, apt and illuminating? Since this is a question about powers and possibilities that is, moreover, couched in metaphorical language, answering it is not a straightforward matter of empirical observation or measurement alone.[32] But while observation alone cannot settle the matter, we might nonetheless draw reflectively, normatively, and critically on developmental psychology and, roughly, on the theory of ego formation for some help. Here it is useful to turn to an account of ego formation that is both pre-Freudian and directly concerned with how awareness of the Kantian moral law as making an authoritative normative demand might arise in the course of subject development.

In his 1795 essay "On the Concept of Punishment," Friedrich Hölderlin takes up exactly this topic. As Thomas Pfau usefully puts it, "with the concept of punishment as a *ratio cognoscendi* of a primordeal order where 'freedom and necessity' seem to have converged, Hölderlin implicitly introduces a temporal marker into Kant's conceptual system."[33] That is, in describing the experience of punishment *as punishment*, Hölderlin seeks to trace how explicit awareness of the normative authority of the moral law arises within the life of a living human subject, *not* as a matter of derivation *from* experience alone, but rather through a dawning exercise of one's distinctive powers as a subject. Here Hölderlin's first move is to argue that what is evil or wrong cannot simply be defined as behavior that *is* punished, that is, as behavior that results in suffering. If that were the case, then, as Hölderlin puts it, "I [would] also deduce an evil will," that is, take myself to have done something wrong, something that ought not to have been willed, "from any other resistance" or experience of suffering. In that case, "all suffering [would count as] punishment."[34] That, surely, is not right. Being caught in the rain and catching a chill may be the result of imprudent action, and it may involve suffering, but it is not a case of being punished by the weather for a moral transgression, nor is stumbling over a crack in the sidewalk and bruising one's knee. What, then, must be added to the experiences of suffering and resistance to one's will in order to have an experience *of punishment?* Here Hölderlin writes, in a fragment that breaks off:

---

32 This is, of course, the truth of the claim that *ought* cannot be derived from *is*, as long as *is* claims are restricted to what is straightforwardly empirically observable or measurable.

33 Pfau 1987, p. 16.

34 Hölderlin 1987b, p. 36.

> To this it may be answered that, insofar as one considers oneself punished, one necessarily implies the transgression of the law within oneself; that in punishment, insofar as one considers it punishment, necessarily [...].[35]

Thus what is crucial is that one considers oneself to be punished, that is, to have done something that merits suffering as imposed by another agent who is enforcing normative demands appropriately. (The relevant suffering can include things like being required to sit on the stairs and think about what one has done; it need not involve only physical pain.) One must grasp that one is suffering appropriately in virtue of having done something one ought not to have done, according to the standards of another agent who is authoritative for one.[36] This grasp has an empirical component—one must have done something and met a reversal—but its content is not entirely empirical, insofar as it includes an emerging awareness of what one ought not to have done in virtue of appropriate standards. Pointedly, this awareness of being punished according to appropriate standards arises in the course of the experience of punishment and reflection on it. Awareness of standards of appropriateness is more an aspect of this reflective experience in a context than it is something already formed and brought to that experience from outside it.

It is possible for there to be human beings, perhaps sociopaths, who do not have such experiences of having been punished appropriately. It is also important to note that having such an experience is not required for being a subject, in the sense of someone who has and is aware of having at least a spatial and subjective point of view on things, including having preferences.[37] Nor is it necessary for being a person, in the sense of someone who is able and entitled under law to make transactions, bring lawsuits, and so forth (perhaps doing so in a pathological spirit). *Subject* and *person* are role- or status-concepts, and it is possible, though happily not normal, to play the relevant roles or to acquire the relevant status without, as it were, developing a conscience. Someone who experienced sufferings and reversals, including corrections by others, but without developing a sense of sometimes having been appropriately punished,

---

35 Hölderlin 1987b, p. 36.

36 This is the truth (consider: 'meaning') of Williams's claim that shame is prior to guilt, since the explicit articulation of what the standard of appropriateness is will come *after* the experience of simply having transgressed an authoritative other's standard.

37 Hence Christine Korsgaard, in taking commitment to the normative authority of the categorical imperative to be part of constituting oneself as a subject, builds too much into a kind of individual faculty psychology and too little into what emerges in interaction with other subjects. See Korsgaard 2009. But then Korsgaard is right that powers inherent to the subject must also be brought into play in this interaction in order to yield this commitment.

would be, as it were, aspect-blind to the existence of an authoritative normative order, blind to the sense that there are things that simply ought not and ought to be done. Though pathologies of development that issue in forms of moral aspect-blindness are possible, they too are happily not normal.

There is, moreover, room for radically significant variation both in behaviors that are punished and in senses of standards of appropriate punishment that may emerge from them. What gets punished and in what ways is in fact significantly different in different cultures. At least in the first instance, then, the experience of punishment *as* suffering *in virtue of* having violated appropriate normative standards need not, and sometimes does not, involve any sense of having violated any single distinctively moral law, let alone a Kantian one.[38] Children and others are in fact punished for all sorts of things, including varieties of rudeness, uncleanliness, acting out, clumsiness, lack of self-control, and so forth, sometimes inappropriately and unjustly so. As a result they sometimes develop standards of appropriate punishment, standards that they may then enforce and transmit, that are in fact themselves unjust and inappropriate. Across cultures and differences in socio-cultural circumstances, what gets punished, what is experienced as punishment, and what is taken to be appropriate punishment may vary widely.

Is there then any reason to think that reference to a single moral law along Kantian lines, such as the formula of respect for persons, could figure and should figure nonetheless in any experiences of punishment as punishment and any developments of conscience? Perhaps the most obvious and straightforward answer to this question is "No; morality in its distinctively Kantian form does have its distinctive socio-cultural circumstances and settings." But while it is true that there are primitive conditions of life, say, where the development of Kantian conscience may not be a reasonable part of the development of conscience as such, this is much less likely to be the case in complex societies, where one is subject to correction and reversal throughout the course of one's life, from infancy onwards, from a variety of authoritative others who themselves have distinct subjective points of view and commitments. Suppose, then, that within a setting of significant interaction with a wide range of diverse others, one simply stood imaginatively, as it were, on the normative authority of only a few, readily identifiable distinct others (perhaps one's parents, perhaps members of a dominant social caste). That is, one experiences punishment as punishment and one develops a sense of appropriate normative standards for it and for

38 This, too, is an insight that Williams has registered, in noting the existence and force for us of highly pluralized 'ethical' demands.

action in general, *only* insofar as these standards are instituted and maintained by members of group G whom one respects and whom feels one must respect. If members of groups H, I, J and so on that one encounters turn out to have different normative standards for appropriate behavior and punishment, then they are simply not to be taken seriously. Perhaps they are members of a dominated and dependent group that one cannot respect or members of a dominant group whom one fears but with whom one does not identify, or perhaps they are simply other, tribally, racially, sexually, economically, or whatever. What if one simply stopped there, in attachment to only the normative standards of distinct group G? This, too, is surely possible and, sadly, often enough actual. Is there any reason to think that this stance is also pathological or a form of normative aspect-blindness?

This question has no ready answer. Attempting to answer it is complicated by the fact that different individuals with different socio-cultural backgrounds may use the same abstract, more or less Kantian language of rights, duties, and respect for persons, but differ dramatically in how they assess individual cases on the ground. What you call telling someone a hard truth out of respect, I call inconsiderate cruelty; or what you call encouragement to develop one's talents and specific forms of self-respect I call indulgence and pampering. Mistakes and errors on all sides are possible, including, and perhaps especially, in describing and judging one's own conduct. What counts as respect for persons is itself a subject for open, imaginative, explorative inquiry and moral conversation in an ongoing way, even where the value of respect for persons is abstractly shared.[39]

But should respect for persons, however contested and evolving its criteria be, figure centrally in the development of any form of conscience arising out of the experience of punishment as punishment? This is a normative, not a factual question. Again, in fact not everyone develops a conscience of this kind, whether out of significant rudeness of circumstances, radical tribalism, or social pathology (as a Kantian view would see it). But it is a normative question that admits of some argument, even if not a priori proof—argument that points toward answering "Yes." For what would it be like simply to live, without further normative reflection, according to the normative standards of (let us say) one's own group G, thence ignoring, dismissing, or violently repudiating the claims

39 This is one way of taking the point of Stanley Cavell's continuing emphasis on the importance of mutually explorative moral conversation rather than 'standing on' one's accomplished theory of right and duty. Leading a moral life is not simply and straightforwardly a matter of applying theory to practice deductively. For Cavell's thought, see part three of Cavell 1979, and also Bates 2003, who traces the theme of moral conversation throughout Cavell's career.

to respect of members of groups, H, I, J... with their local practices? It would be, arguably, to stop short in reflection, to do a kind of violence to oneself in restricting one's concept of conscience to the standards of only a distinct and recognizable few whose authority one had internalized, thereby cutting oneself off from significant practical interaction with members of other groups. It would amount, as it were, to freezing oneself as a G, acknowledging only the local normative standards of G-ness. (And of course this goes, conversely, for freezing oneself in H-ness, I-ness, J-ness and so on too.)

This may sound all too open, tolerant, and pluralist, as though we were to embrace the normative standards of any group whatsoever. But that is not right. Rather, from within the development of conscience in initial settings of authoritative others whose standards one internalizes, it is both possible and worthwhile to develop a generalized conception of the value of respect for human beings as ends in themselves, that is, as both beings who can set their own ends and who are capable themselves of normative reflection and the development of a generalized conception of conscience. Such a conception requires a kind of tolerance and pluralism, but a kind that is compatible with mutual criticism and sharp disagreement.[40]

It is therefore both possible and normatively apt that, as Kant puts it, we will "*with time* transform the rude natural predisposition [Anlage] to make moral distinctions into determinate practical principles" on which we may act more effectively, so as to "transform a *pathologically* compelled agreement to form society into a *moral* whole."[41] "Steps from crudity toward culture"[42] are possible, and moral perfection is to be sought and hoped for "from nowhere else but education."[43] There are, however, no guarantees. We *may* make "the transition from the go-cart of instinct to the guidance of reason—in a word, from the guardianship of nature into the condition of freedom,"[44] but that depends on us, specifically, on how we exercise our powers of reflection and commitment first in situations of dependence on a few authoritative others and later in relation to more wide-ranging encounters.

Apt acknowledgments of a Kantian moral law will, moreover, not be available or fruitful at every historical moment. There are, again, circumstances both

40 For an excellent articulation and defense of the kind of toleration that we ought to practise and that involves and requires mutual engagement and criticism rather than mutual indifference, see Oberdiek 2001.

41 Kant 2007c, p. 111; first emphasis added.

42 Kant 2007c, p. 111.

43 Kant 1997b, p. 221.

44 Kant 2007b, 8:115, p. 168.

prior to morality in the Kantian sense (prelinguistic infants; rude conditions of nomadic life) and beyond morality (extreme emergencies, involving treacheries and evils to be confronted, where no option is innocent). Bernard Williams is thus right to some extent when he observes that "the drive toward a *rationalistic conception of rationality* comes ... from social features of the modern world, which impose on personal deliberation and on the idea of practical reason itself a model drawn from a particular understanding of public rationality."[45] That is, it is true that commitment to the relevance of deliberation guided by universal and impersonal principles has social conditions of emergence, though those conditions, contra Williams, evidently obtained in both fifth century B.C.E. Athens and in first century C.E. Rome and Palestine, among other places. These conditions arise in various shapes more or less whenever human settlements have become large enough to require significant political and economic role differentiation and regular interactions among occupiers of different roles.

III

It is, then, at least possible to acknowledge from within reflection the requirements of a moral law of respect for human beings, and it is perhaps more likely for such acknowledgment to be arrived at by more people the less primitive and the more cosmopolitan the circumstances of life are, where one encounters and internalizes the normative authority of a wide range of others. But if such acknowledgment does not necessarily take place even in such circumstances, then what is the *normative point* of it? Why be moral (by more or less Kantian lights)?

In Part III of *A Theory of Justice*, John Rawls undertakes to show that a just scheme of cooperation, if instituted, would be stable, even in the face of widespread possibilities of free riding. He appeals to existing "relations of friendship and mutual trust, and the public knowledge of a common ... sense of justice" that is "normally effective" in shaping the actions of most adult subjects of the scheme.[46] Those growing up under such a scheme initially lack such relations and such a sense of justice, but they can be expected to develop it, insofar as it is "given that family institutions are just and that the parents love the child and manifestly express their love by caring for his good" and given that the child then "develops ties of friendly feeling and trust toward others in the association

45 Williams 1985, p. 18.
46 Rawls 1999, p. 435.

as they with evident intention comply with their duties and obligations, and live up to the ideals of their station."[47] One might put this point by saying that the child's developing sense of justice *sanctifies* relations, roles, and values that are already in place and in which he is coming to participate more self-consciously, just as a declaration of commitment in a marriage ceremony may sanctify a form of mutuality in place that demands continuation and acknowledgment.

Rawls denies that the availability and even the naturalness under favorable enough conditions of acknowledgment of the moral law constitute a justification of the obligation to be moral. As he puts it, "these considerations do not determine the ... acknowledgment of principles."[48] This is in part because he is thinking of justification primarily as a matter of the availability of a set of fully articulated public considerations that are sufficient to determine once and for all a choice of value schemes from among all plausible available alternatives.[49] But that is not the only conception of justification that is relevant to moral theory, and Rawls also observes that "these considerations ... confirm" that acknowledgment and that it may "happen that the superiority of a particular view (among those currently known) is the result, perhaps the unexpected result, of [a] newly observed consensus" that is subsequent to critical and comparative reflection. [50] From this point of view, the question of how we come developmentally in favorable enough circumstances to acknowledge the moral law is not in the end separable from the question of what reasons we have to acknowledge it, even if the value of such acknowledgment must also always be tested in reflection. We cannot have, but happily do not need, a justification for being moral "from nowhere," apart from our location within a set of developing circumstances of life.

47 Rawls 1999, p. 429.

48 Rawls 1999, p. 508.

49 "Justification is argument addressed to those who disagree with us, or to ourselves when we are of two minds. It presumes a clash of views between persons or within one person, and seeks to convince others, or ourselves, of the reasonableness of the principles upon which our claims and judgments are founded. Being designed to reconcile by reason, justification proceeds from what all parties to the discussion hold in common. Ideally, to justify a conception of justice to someone is to give him a proof of its principles from premises that we both accept, these principles having in turn consequences that match our considered judgments. Thus mere proof is not justification. A proof simply displays logical relations between propositions. But proofs become justification once the starting points are mutually recognized, or the conclusions so comprehensive and compelling as to persuade us of the soundness of the conception expressed by their premises." (Rawls 1999, p. 508).

50 Rawls 1999, pp. 508, 510.

The formula of the moral law, in any of its Kantian versions, will not itself be fixed by a priori reason, but instead held in view as a kind of summary of a commitment that one has come to find unavoidable within one's life, initially rooted in the experience of punishment, of a being able to traffic in obligatory, non-instrumental norms.[51] Nor will epistemic certainty be available. Commitment to the moral law, motivated by critical discernment, will remain an actively maintained stance, not the product of any discernment-neutral theory modeled on the sciences.[52] (As Wittgenstein once wrote, "I must plunge into the water of doubt again and again."[53]) But the normative value of maintaining this stance is there to be discerned from within a wide-range of adequate circumstances, including virtually all non-emergency circumstances within modern settled life and involving significant interactions across group boundaries. To know the moral law is to acknowledge it from within what, who, and where one is.[54]

Bibliography

Ameriks, Karl (2012): *Kant's Elliptical Path*. Oxford: Oxford University Press.
Bates, Stanley (2003): "Cavell on Ethics". In: Eldridge (2003).
Budick, Sanford (2010): *Kant and Milton*. Cambridge, MA: Harvard University Press.
Cavell, Stanley (1979): *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. New York: Oxford University Press.
Eldridge, Richard (1989): *On Moral Personhood: Philosophy, Literature, Criticism, and Self-Understanding*. Chicago: University of Chicago Press.

51 Compare Eldridge 1989, pp. 47 – 50, 63 – 67. 181 – 88.

52 Here the view should be contrasted with the constitutivist view urged by Christine Korsgaard, according to which "what it is to be a person, or a rational agent, is just to be engaged in the activity of constantly making yourself into a person" where "the categorical imperative," as "constitutive of action," is unconditionally normatively binding on this process of making. (Korsgaard 2009, pp. 42, 52) Against this, Karl Ameriks is surely right to point out that "in ordinary language one would not be called irrational simply because one has not achieved or approached a so-called rationally maximal state" (Ameriks 2012, p. 154), so that it will be impossible to ground commitment to the binding authority of the categorical imperative solely in the structure of rational agency.

53 Wittgenstein 1931 p. 119.

54 Robert Louden and Beatrix Himmelmann pressed me to develop the moral psychology presented here in less exclusively Kantian terms and to be explicit about the relevant concept of justification. Hans Oberdiek and Krista Thomason read and commented on an earlier draft of this essay. Earlier versions of this essay were presented to audiences at a joint University of Sydney-University of New South Wales conference on Nature and Culture in German Romanticism and Idealism and at the philosophy department of the University of Western Australia, in each of which settings I received useful responses. I am grateful to all these individuals and audiences.

Eldridge, Richard (Ed.) (2003): *Stanley Cavell*. Cambridge: Cambridge University Press.
Freeman, Samuel (2012): "Why Be Good?". In: *The New York Review of Books*, 59, 7, pp. 52–54.
Hölderlin, Friedrich (1987a): *Essays and Letters on Theory*, ed. Thomas Pfau. Albany: Suny Press.
Hölderlin, Friedrich (1987b): "On the Concept of Punishment", trans. Thomas Pfau. In: Hölderlin (1987a), pp. 35–36.
Kant, Immanuel (1959): *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck. Indianapolis: Bobbs-Merrill.
Kant, Immanuel (1996a): *Practical Writings*, ed. Mary J. Gregor. Cambridge: Cambridge University Press.
Kant, Immanuel (1996b): *Critique of Practical Reason*, trans. Mary J. Gregor. In: Kant (1996a).
Kant, Immanuel (1996c): *The Metaphysics of Morals*, trans. Mary J. Gregor. In: Kant (1996a).
Kant, Immanuel (1996d): *Religion and Rational Theology*, eds. Allen W. Wood and Giorge di Giovanni. Cambridge: Cambridge University Press.
Kant, Immanuel (1996e): *Religion with the Boundaries of Mere Reason*, trans. George di Giovanni. In: Kant (1996d), pp. 55–215.
Kant, Immanuel (1997a): *Lectures on Ethics*, eds. Peter Heath and J. B. Schneewind. Cambridge: Cambridge University Press.
Kant, Immanuel (1997b): "Moral Philosophy: Collins's Lecture Notes", trans. Peter Heath. In: Kant (2007a), pp. 37–222.
Kant, Immanuel (2007a): *Anthropology, History, and Education*, eds. Günter Zoller and Robert B. Louden. Cambridge: Cambridge University Press.
Kant, Immanuel (2007b): "Conjectural Beginning of Human History", trans. Allen W. Wood. In: Kant (2007a), pp. 163–175.
Kant, Immanuel (2007c): "Idea for a Universal History with a Cosmopolitan Aim", trans. Allen W. Wood. In: Kant (2007a), pp. 108–120.
Korsgaard, Christine (2009): *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.
Hegel, G. W. F. (1991): *Elements of the Philosophy of Right*, ed. Allen W. Wood, trans. H. B. Nisbet. Cambridge: Cambridge University Press.
MacIntyre, Aladair (1981): *After Virtue*. Notre Dame, Indiana: University of Notre Dame Press.
Oberdiek, Hans (2001): *Tolerance: Between Forbearance and Acceptance*. Boston: Rowman & Littlefield.
Parfit, Derek (2011): *On What Matters*, 2 vols. Oxford: Oxford University Press.
Pfau, Thomas (1987): "Critical Introduction". In: Hölderlin (1987a), pp. 1–29.
Pinkard, Terry (2002): *German Philosophy 1760–1860: The Legacy of Idealism*. Cambridge: Cambridge University Press.
Rawls, John (1999): *A Theory of Justice*, rev. ed. Cambridge, MA: Harvard University Press.
Williams, Bernard (1985): *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
Williams, Bernard (1993): *Shame and Necessity*. Berkeley: University of California Press.
Williams, Bernard (2008): *Philosophy as a Humanistic Discipline*. Princeton: Princeton University Press.
Wittgenstein, Ludwig (1931): "Remarks on Frazer's *The Golden Bough*". In: Wittgenstein (1993), pp. 118–155.

Wittgenstein, Ludwig (1993): *Philosophical Occasions, 1912 – 1951*, eds. James Klagge and Alfred Nordmann. Indianapolis: Hackett.

Erik Lundestad

"Why Be Moral?" Pragmatism's Attempt to Dismiss the Issue

Pragmatists hold that the question, "Why be moral?" both can and ought to be dismissed. As they see it, this question only appears interesting for those of us who, mistakenly, assume that morality forms a sphere distinct from that of prudence. Given this misconception, moral acts will be seen as having a specific aim, different from that of prudential acts. Thus, the question will appear: Why do we act *morally?* What are *moral* acts aimed at? Pragmatists, however, dismiss the notion of there being such a gap between morality and prudence. This distinction, they argue, does not have any precursor in practice, in the manner in which we act. Since, according to pragmatism, everything that we do is aimed at a good, the question, "Why be moral?" just isn't very interesting. But is this pragmatic dismissal of the "Why be moral?" issue convincing? Can the question, "Why be moral?" really be dismissed in this manner? Since pragmatism contains a wide range of approaches, I will in the following be discussing these issues in relation to what I myself consider the most promising of these, namely that of John Dewey.

I

Like other pragmatists, John Dewey takes the interest surrounding the "Why be moral?" question to result from the notion that there is a gap between that sphere in which we act in order to obtain something that we consider good and that sphere in which we act on the basis of other, specifically *moral* reasons. Like other pragmatists, moreover, Dewey believes that everything we do is done in order to obtain a good. As he sees it, therefore, there is no such gap. This, however, is not something which it is sufficient merely to *claim*. What must be done, rather, is to *show* that the notion of such a gap is superfluous, i. e. that there is no need to appeal to it in order to make sense of what we do. Dewey's main philosophical strategy, therefore, is that of *reconstruction*. This implies that what are perceived as distinctively "moral" acts, must be reconstructed – or "re-described", as the neo-pragmatist Richard Rorty would later call it – so that they too may be seen as something which is done in order to obtain a good. The test as to whether the "Why be moral?" issue can rightly be dismissed, we are thus in po-

sition to state, is whether all that we do *can* be reconstructed pragmatically – as done in order to obtain a good.

It is in accordance with the philosophical strategy which has been outlined above that John Dewey, in his 1908 *Ethics*, argues that concepts such as 'duty', 'right' and 'obligation' may be seen as resulting from our need to make up, or compensate for, temporary failures in the way that we adjust to our environment (Dewey 1978, p. 310). Dewey's point is that since we do not always adjust successfully to our surroundings, we sometimes need to be told that we have a duty or obligation to do this or that. According to Dewey, the purpose of notions such as these is therefore to correct our behavior so that what we do in fact is aimed at something good.

In well-known works such as *Human Nature and Conduct* (1922) and *The Quest for Certainty* (1929) this same approach manifests itself in the manner in which Dewey emphasizes the establishing of *habits*, that is of how children learn to anticipate how other people will react to what they do, and how, in turn, these reactions are internalized by the child. The result of successfully having established habits, Dewey may therefore emphasize, is that in a given instance there is no gap between the child's own conception of the good and what truly *is* good.

It should be stressed that Dewey does not imply that we will ever reach a stage in which concepts such as 'duty', 'right' or 'obligation' may in fact be left behind. Dewey, to the contrary, explicitly criticizes notions of a 'final end' (Dewey 2002, p. 174 – 175). What he points out is that these concepts have a specific, empirical *purpose or end*, namely to ensure that our notions of the good do not lose touch with reality – with what in fact *is* good.

There are several problems related to a pragmatic approach of the kind now considered. The one problem I will be concerned with in what follows is that it runs the risk of conflating what *is good* with that which a contingent society at a given time actually *takes to be good*. Those familiar with pragmatism will identify this problem as a parallel to the (better known) problem of whether pragmatism mistakenly conflates what *is true* with what we in fact *take to be true*.

Some (neo-) pragmatists, most notably perhaps, Richard Rorty, would be more than happy with such a conflation. As Rorty sees it, to be moral simply *is* to have internalized a specific set of values and norms. If I say, making use of first person, "We" or "People of our sort" don't do this, Rorty would argue, I am reporting a norm. If, on the other hand, I stand back from my community, and report, in the third person: "They" or "People of their sort" don't do this, then I am reporting a fact. For Rorty the source of the norm is the internalization of the fact, or vice versa, the source of the fact is the externalization of the norm (Rorty 2007, p. 196 – 197).

The reason that we act morally according to Rorty is thus simply that we have made certain values and norms our own. Rorty cites Christine Korsgaard who claims that a successful answer to the question, "Why be moral?" "must show that sometimes doing the wrong thing is as bad or worse than death" – and who goes on to note that "the only thing that could be as bad or worse than death is something that for us amounts to death – not being ourselves any more" (Rorty 2007, p. 197). Rorty agrees to this, but regrets that Korsgaard takes it to imply that there must be something other or more going on than that we have internalized a certain set of norms.

Rorty's view is highly controversial, and for good reasons. In a much discussed scene from Mark Twain's novel, *The Adventures of Huckleberry Finn* (which Rorty himself refers to), Huck Finn eventually decides to help Jim, the runaway slave, escape rather than turn him in. He thus acts against both the values and norms of the society of which he is part and his own conscience. As Rorty sees it, Huck Finn is, by doing this, simply taking on another moral identity (Rorty 2007, p. 197). Since, for Rorty, there is no such thing as morality in distinction to our actual, historically constituted, and therefore contingent, *concepts* of morality, there is no non-circular way for us to argue that helping Jim escape is the right thing for Huck Finn to do.

Dewey would agree with Rorty that being moral is to have made certain values and norms one's own. He would not agree, however, that the values and norms we adhere to are merely contingent. Unlike Rorty, Dewey would insist that Huck Finn, when deciding to help Jim escape rather than report him, is doing what he ought to do in a sense of "ought" which is not merely relative to a contingent set of values and norms. But, and this is the crucial issue for a pragmatist approach to morality, what sense of "ought" can this possibly be? Mustn't pragmatists, given that they distance themselves from the notion of a distinctive moral realm, also distance themselves from the notion of there being any other sense of "ought" than the one we are in position to identify on the basis of our actual values and norms? May it not be plausibly argued therefore, that Rorty's conflation is what pragmatism, in effect, must come down to?

To address this issue, let's return to Dewey's 1908 *Ethics*. In this work, Dewey proposes that to deliberate can be seen as performing a "dramatic rehearsal" (Dewey 1978, p. 292). What this means is that when we deliberate we let possible courses of action unfold themselves in our imagination so that we are able to foresee the effects they will have in real life. This notion of a "dramatic rehearsal" provides the basis for the later, full-fledged pragmatic approach Dewey is to take in works such as *Human Nature and Conduct* and *The Quest for Certainty*. In these works he suggests that "we regard our direct and original experiences of

things liked and enjoyed as only possibilities of values to be achieved, that enjoyment becomes a value when we discover the relation upon which its presence depends" (Dewey 1960, p. 259).

What Dewey here is saying is that when we deliberate over what we ought to do, we always start out with our own, actual beliefs of something as valuable or good. Rather than identifying that which – objectively speaking – *is* valuable or good with these beliefs, however, Dewey suggests that we should approach these beliefs as potential candidates for what is valuable or good in an objective fashion. To settle whether these actual beliefs truly are valuable or good, therefore, we need to *test them*. What Dewey is suggesting, in other words, is thus that our beliefs regarding the valuable or good should be seen as *hypotheses open for testing*. Here is Dewey:

> Any belief as such is tentative, hypothetical, it is not just to be acted upon, but is to be framed with reference to its office as a guide to action. Consequently, it should be the last thing in the world to be picked up casually and then clung on to rigidly. When it is apprehended as a tool and only a tool, an instrumentality of direction, the same scrupulous attention will go to its formation as now goes into the making of instruments of precision in technical fields. [...] A moral law, like a law in physics, is not something to swear by and stick to at all hazards; it is a formula of the way to respond when specified conditions present themselves (Dewey 1960, p. 278).

This passage captures the distinctive features of Dewey's pragmatism, or "instrumentalism" as he himself was going to call it. Beliefs, he states, should be seen as *tools*, as *guides to action*.

It must be emphasized that Dewey is *not* suggesting that beliefs are "subjective" in the sense that it is "up to me" to decide which belief best satisfies my own interests, desires or needs. Part of the reason why the need for a distinct moral sphere arose in the first place, Dewey believes, is because of the misguided assumption that we in fact are self-interested individuals. Once this assumption has been made, it becomes necessary to introduce the notion of morality as something "over" or "above" the empirical sphere in order to explain why not all of our actions are self-interested. This, Dewey believes, is the basis of the moral philosophy of Kant. According to Dewey, Kant "carries to the logical extreme [the misguided] notion of the opposition between all the values which satisfy desires and the true moral good" (Dewey 1985, p. 220). This opposition, however, does not merely isolate morality from our doings at large, but also serves to isolate the empirical sciences from morality:

> The narrow scope which moralists often give to morals, their isolation of some conduct as virtuous and vicious from other large ranges of conduct, those having to do with health and vigor, business, education, with all the affairs in which desires and affection are implicated,

> is perpetuated by this habit of exclusion of the subject-matter of natural science from a role in formation of moral standards and ideals. The same attitude operates in the other direction to keep natural science a technical specialty, and it works unconsciously to encourage its use exclusively in regions where it can be turned to personal and class advantage, as in war and trade (Dewey 1960, p. 274).

Dewey's own position, in contrast, is that any kind of deliberation (whether it is concerned with morality, science or everyday life) takes place against a background of already established social practices. Why, say, do we eat bread rather than porridge for breakfast? Why do we greet people as we do, rather than in some other way? There are no ready answers for such questions. It is just the way we do things. If, however, things do not turn out as we are used to, if say, a problem occurs, we experience doubt as to how to proceed. It is this *experience of doubt*, Charles Sanders Peirce, the founder of pragmatism, would say, which "sets the problem" (Peirce 1955, p. 11). According to pragmatism therefore, any attempt to 'secure' inquiry by ridding us of all unwarranted assumptions, will have to be seen as doomed from the start.

What has been said also implies that for any belief to put an end to our experience of doubt, it must be able to "fit in" with other parts of our experience, that is, with other facts that are known to us, with other beliefs that we consider valid, and so forth. And again, the fact that a belief may be able to end our doubt *here and now*, does not imply that it may not, in the future, give rise to other problems, which is why Dewey in the cited passage stresses that beliefs should not be "clung on to rigidly" but be regarded as fallible, or as open for future revision.

Another way of stating this point would be to say that Dewey is advising us to follow the lead of the modern, empirical sciences. Within these sciences, Dewey believes, we do not only cooperate in reaching a common goal, but it is commonplace for everyone to treat their beliefs as fallible. Scientists are always prepared, ideally at least, to revise their theories. It is this very attitude, Dewey believes, that should provide the basis for how we approach all areas of life.

So why be moral? According to what has been said, there isn't much to be said about this. Since, according to Dewey's pragmatism, there are no distinct moral acts, there is no specific realm of morality "over" or "beyond" the realm of prudence; to act morally is to act in order to realize a good. This, Dewey would argue, is what we do, for the most part at least, when we raise children, advise students, help customers, do research, take part in political decisions and so forth. It may turn out, of course, that we in our teaching, say, have made use of a method which has some drawbacks, or that we, in our med-

ical practice, have made use of a medical procedure which, as it turns out, has certain less advantageous effects. Life does not come with any guarantees. The point is merely that when such things happen, we should be prepared to revise our methods and procedures. All we *can* do is to adjust our activities as best we can, as we go along. Since, according to what has been said, we are all social beings concerned with the well-being of others, there just isn't any need for a notion of morality as something "over" or "above" this.

II

Not everyone will agree, of course, that all of our issues can be settled empirically. According to Jürgen Habermas it is a mistake to assume that we can distinguish what is *rightfully* considered valuable or good from what we merely *take* to be valuable or good by way of testing our beliefs, in the manner of Dewey's pragmatism. What enables us to draw this distinction according to Habermas is not, as Dewey mistakenly thinks, the existence of an empirical reality which somehow "allows" us to perform some acts successfully whereas it "denies" us the possibility of performing other acts. It is the existence of people, rather, with values and norms that are *different from our own*. The only reason Dewey fails to recognize this, Habermas believes, is because he presupposes from the outset that we have common values and norms. Dewey approaches morality "from the perspective of community members concerned with their common good" (Habermas 2005, p. 234). The moral agreement in which Dewey's pragmatic approach is intended to *result* is in fact *presupposed already from the outset!*

The basic reason Habermas takes Dewey's pragmatism to be flawed is thus that it is based on what he – Habermas – refers to as a "vertical We-perspective." Within this perspective all members of society "[top to bottom] can identify everyone else as a member of *the same* cooperative community" (Habermas 2005, p. 234). Given a vertical We-perspective, therefore, the issue as to which values and norms we ought to adhere to is already decided. The only issue at stake is *how* these common values and norms best can be accommodated. This, to be sure, *is* an empirical issue. The problem, however, is that the "vertical We-perspective" cannot form the basis in "discourses that cross communal boundaries." The reason for this, he claims, is that in *such* discourses "the participants take on a first-person-plural perspective that is not vertically directed at all members top to bottom, but *horizontally* at the mutual inclusion of the other." (Habermas 2005, p. 234) It is thus only on the basis of a distinctive "*horizontal* We-perspective" that we are able to account for morality, according to Habermas. He then goes on to point out that it was another classical American pragmatist,

namely George Herbert Mead, who emphasized how such a point of view is "generated by all participants symmetrically and reciprocally taking on each other's perspectives" (Habermas 2005, p. 234).

The case of Huck Finn may thus be taken to illustrate what Habermas takes to be the problem. The problem, it seems, is that as long as Dewey's pragmatism leaves out the so-called "horizontal" We-perspective altogether, it will repeat Huck Finn's mistake of identifying that which ought to be done with that which his own society *believes* ought to be done. If Huck Finn had been a pragmatist of Dewey's kind, it would therefore seem, he would test whether it is helping Jim escape or reporting him which contributes to the larger social good. The depressing answer that he would have come to, given the values and norms of the society of which he is part, would thus be that the morally responsible thing to do would be to report Jim.

It would seem therefore, that Dewey's pragmatic attempt to dismiss the "Why be moral?" question by way of undermining the gap between morality and prudence, results in precisely that kind of conflation which is defended by Rorty. This conflation, most of us would argue, is not satisfactory since it rules out the possibility of a (non-circular) way of approaching the issue as to which values and norms we *ought* to adhere to. In the last instance, it would seem, it simply *replaces* morality with prudence. The lesson to be learned, it therefore seems, is that Kant and Habermas are right in claiming that the notion of a distinct moral sphere must be upheld.

The presented criticism may appear devastating. There is nevertheless a response. As we have seen, one of the main points of Dewey's pragmatism is that everything we do is done in order to realize some desire or need. For a pragmatist, this is simply what it *means* to act. As Dewey points out in *Human Nature and Conduct*, to be a living being *is* to have specific desires and needs to satisfy. To transcend these would thus be to transcend life itself.

But not only is it in fact impossible to overcome the very sphere of our desires and needs as such. It doesn't even make sense to uphold doing so as an ideal. The reason for this is that it is only *because* we have certain desires and needs that we are able to experience something *as* something, i.e. as, cold, dry, satisfying, missing, beautiful or exact. Take away our desires and needs, therefore, and you take away the possibility of experiencing something as something altogether. The notion of morality as constituting a distinct sphere "over" that of prudence therefore, doesn't make sense. If such a sphere existed it would, according to the point now made, necessarily have to be empty. This line of criticism may not only be applied to Kant, but also, as Albrecht Wellmer has shown, to Habermas. As Wellmer points out, Habermas' very notion of an "ideal communication community" must in itself be seen as misguided, since even to imagine

such a community is to imagine a situation in which there would be nothing left to communicate (Wellmer 1998).

It would seem therefore, that we are confronted with a dilemma. From the point of view of Habermas and others, it may be argued that the result of any attempt to do *without* the notion of a distinct moral sphere will be that we miss out on morality altogether. On the other hand, a pragmatist such as Dewey may, with just as much right, argue that any notion of such a distinct sphere will have to be empty since to distance ourselves from our desires and needs as such is to distance ourselves from the very possibility to experience something as something.

It is, I suspect, this very dilemma that Rorty's provocative form of neo-pragmatism is trading on. Even though most of us will find Rorty's conflation of what is moral with what we take to be moral highly unsatisfactory, Rorty may still defend his position indirectly, i. e. by pointing out that there is no alternative, since the notion of a distinct moral sphere may be dismissed as absurd. Since, however, Rorty merely settles for one horn of the dilemma, this is a strategy which goes both ways. His opponents, such as Habermas, may therefore, with just as much right, defend their position indirectly, i. e. by pointing out that Rorty misses out on morality altogether.

The implication to be drawn from this is that if pragmatism is to be successful in its dismissal of the "Why be moral?" question, then it will have to come up with a satisfying solution to this dilemma. It will have to come up with an approach that maintains the autonomy of the moral sphere (so that it cannot be (rightfully) accused of missing out on morality altogether), but, without cutting the relation to desires and needs completely (so that it cannot be (rightfully) accused of presenting us with an empty notion of morality). Can this be done?

III

It is well established that Dewey came to revise his own approach to morality. This revision takes place with "Three Independent Factors in Morals", a lecture he gave to the French Philosophical Society in 1930. Here Dewey for the first time distinguishes between three *independent* factors in morality. He thereby distances himself from the approach taken in the original 1908 *Ethics*, where he had argued that notions such as 'duty', 'right' and 'obligation' can and ought to be analyzed on the basis of the notion of the 'good' (Pappas 2008, p. 94). This revision, I would like to argue, may be seen as resulting from an attempt to solve the dilemma which was presented above. To illustrate how Dewey *now* ap-

proaches the issue, consider the following passage from the revised 1932 edition of his *Ethics*, describing the relation between a parent and a child:

> A child may be subject to demands from a parent which express nothing but the arbitrary wish of the latter, plus a power to make the child suffer if he does not conform. But the claims and demands to which the child is subject *need* not proceed from arbitrary will; they may issue from the very nature of family life in the relation which exists between parent and offspring. Then they do not come to the child as an external and despotic power, but as expressions of a whole to which he himself belongs. He is moved to respond by his affection for his parents, by his respect for their judgment; even when the demand runs contrary to his uppermost desire he still responds to it as something not wholly alien. Because of inherent relationships persons sustain to one another, they are exposed to the expectations of others and to the demands in which these expectations are made manifest (Dewey 1985, p. 218).

Previously, we remember, Dewey had focused rather one-sidedly on how we learn to adapt to the claims of our parents and thus to already existing values and norms. It is, we have seen, precisely this one-sidedness that Habermas takes to undermine Dewey's position. As he sees it, it results in Dewey's assumption that established values and norms are given and thus isolated from criticism. In passages such as the one cited above, however, we find the beginnings of an alternative approach to this issue. Rather than focusing on how the child adapts to already established values and norms, Dewey is now approaching values and norms as something which results from the mutual claims and demands of those who are involved.

In the cited passage, we can see, Dewey is focusing on the distinction between arbitrary and non-arbitrary claims. Some claims, he states, are arbitrary because they are based merely on the parent's own desires and needs. Other claims, however, are non-arbitrary because they are based on what Dewey calls the "very nature" of the relation between parent and child. To the degree that a claim or demand is based on the "very nature" of this relation, he goes on to argue, it will not be viewed as resulting from a force external to the child, but will be seen rather, as "an expression of a whole to which he himself belongs."

This difference between arbitrary and non-arbitrary claims, Dewey then goes on to point out, also manifests itself in the way in which the child reacts vis-à-vis claims. Whereas a child acting on behalf of a claim which it recognizes as arbitrary may do so in order, say, to avoid punishment, a child acting on behalf of a claim that it recognizes as non-arbitrary, may do so out of *respect for the judgment* of the parent. Even if, in this last instance, there may exist a gap between what the child wants to do and what the parent tells it to do, the child may nevertheless choose to do what the parent tells it to do, out of respect for the judg-

ment of the parent. In a situation such as this, we may thus say, the claim is acknowledged as *legitimate* or *right*.

Dewey's point, to be sure, is not to suggest that the parent and the child are equal negotiators. The point, rather, is that not everything that the parent does vis-à-vis the child is in the child's best interest. Rather the child has *its own* claims and demands (even though it may not be able to give expression to them as such). So, just as the parent may raise questions regarding the claims and demands of the child, the child may raise questions regarding the claims and demands of the parent:

> The case is perhaps even clearer if we consider the parent as one who is also subject to claims. These need not be voiced in explicit form by the child; they do not proceed consciously from him. But the parent who is conscientious feels that that they are involved in the *parental relation*. Because of this human relationship, something is owed to the child even though (perhaps even more because) the latter is not able to formulate that claim in any express demand (Dewey 1985, p. 218).

On this basis it may be argued that what Dewey now is offering us, albeit only in outline, is precisely that which seemed lacking in his previous approach. Whereas he previously had focused on how we address problems against a common set of values and norms, he is now introducing the notion of a sphere in which a *plurality* of claims and demands are openly discussed and criticized. He is presenting us, in short, with the notion of a distinct moral sphere.

So, rather than seeing notions such as those of 'duty', 'right' and 'obligation' as resulting from the need to compensate for failures in the way we adjust to the environment, as Dewey had done in the original 1908 edition of the *Ethics*, he now sees these notions as resulting from the relations we have to other people, with distinct claims and demands. So, just as I, according to this view, may have certain obligations or duties towards you, you have certain obligations or duties towards me. Even though, moreover, what you may demand of me may conflict with what I myself *want* to do, I may nevertheless, according to this revised view, recognize your demand as legitimate and therefore come to the conclusion that I ought to act in accordance with it, rather than do what I prefer to do. In cases such as this the right will thus appear as distinct from the good; morality as distinct from prudence. "The Good is that which attracts", Dewey now states, whereas "the Right is that which asserts that we *ought* to be drawn by some object whether we are naturally attracted to it or not" (Dewey 1985, p. 216 – 217).

But does not this imply that Dewey simply has given up on his own, pragmatic position? No, it does not. The point is that even if he is now insisting that the right must be analyzed independently of the good, he nevertheless maintains that there is a relation between the two. According to what has

been said, the right signifies something that a person does not in fact consider his or her own good, but that he or she nevertheless *should* take to be a good. It is this element of 'should', we have seen, that differentiates the notion of the 'right' from that of the 'good'. This, however, does not imply that a gap therefore exists between the two, since what 'should be' is that someone finds a required way of acting good. It may thus be asked whether the conduct that *ought to be performed*, which allegedly has moral authority, *actually contributes to a good* in which the person upon whom the duty is laid will share. According to Dewey, this last question, to be sure, is an empirical one. It may therefore be tested.

So, even though Dewey no longer accounts for moral principles and beliefs on the basis of our dealings with the environment, but rather by way of our confronting a plurality of people with distinct claims and demands, this does not rule out that these principles and beliefs can be tested empirically. Such a testing, he insists, will settle whether what we consider right in fact is good.

On this basis, we can begin to see the outlines of the way in which Dewey intends to solve the previously mentioned dilemma. Whereas the moral and the prudential must be recognized as two (relatively) independent spheres, each with its own set of concepts and logic, one must simultaneously maintain that there must be a "feedback relation" between the two (Hickman 2007). That is, each of the two spheres both can and ought to be evaluated in light of the other.

We have seen that according to Habermas, Dewey's pragmatism presupposes the existence of a common set of values and norms. That is, it presupposes a "Vertical We-perspective". This, Habermas, believes, implies that the testing of beliefs is merely able to decide which *means* best serve the *purposes* or *ends* that these already established values and norms provide us with. The distinct question of which values and norms we *ought to* adhere to is ignored.

Even though Dewey spends a considerable amount of time trying to undermine precisely that clear-cut distinction between means and ends on which Habermas' criticism is based, it must be said that Dewey's works prior to the revision do give the impression that he presupposes a common set of values and norms and that he therefore emphasizes the "vertical" *rather* than the "horizontal" We-perspective. He is thus, we may say, merely settling for one horn of the dilemma.

In his revised stance, on the other hand, Dewey makes it quite clear that to emphasize the one need not rule out the other. In his later work, therefore, Dewey is stressing the relative autonomy of the spheres, and thus the *continuity* of morality and prudence, of ethics and the empirical sciences. He is, as Larry A. Hickman puts it, "envision[ing] scientific technology and communicative action as continuous with one another and as features of a larger inquirential project..." (Hickmann 2007, p. 68 – 69). In this manner, pragmatists of Dewey's kind believe

that it is possible to leave behind the dilemma that seems to beleaguer Habermas. In Hickman's words, it becomes possible for them to argue that Habermas leaves us with "an underlying dualism from which his project continues to suffer" (Hickmann 2007, p. 69).

IV

The cornerstone of Dewey's pragmatism, we have seen, is that (moral) principles and beliefs should not be thought of as belonging to a distinct sphere, but should be revised on the basis of the empirical knowledge that we have—both of ourselves and of the environment in which we live. What does this approach amount to in practice? Do we have any instances to which we can point?

One contemporary instance of precisely this approach may be found in the way in which Cass R. Sunstein, professor of law at Harvard, has argued that we ought to approach problems of public policy. His highly influential approach is presented in a number of books, including *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2008, co-written with Richard H. Thaler), *Simpler: The Future of Government* (2013) and *Why Nudge? The Politics of Libertarian Paternalism* (2014). Even though Sunstein is focusing more on political rather than moral issues his approach is, as we shall see, of a distinctively Deweyan kind.

Between 2009 and 2012, Sunstein served as head of the Office of Information and Regulatory Affairs (OIRA). This office was created in 1980 by the so-called "Paperwork Reduction Act". Under this act, no federal agency in the United States is allowed to make you fill out a form unless OIRA allows it to do so (Sunstein 2013, p. 2). In 1981 President Reagan gave OIRA an even more important role, namely to oversee federal regulation. In *Simpler: The Future of Government* (2013) Sunstein accounts for his time at OIRA, and for what he considers a new approach to the issue of regulation. The following brings forth the basic features of this approach:

> Should we rely entirely on the free market to protect the safety of food? To protect clean air? Even the most market-friendly economists recognize the existence of market failures, which can justify regulation. Suppose that food safety and clean air regulation can save large numbers of lives and do so at a low cost. Or suppose, even, that such regulation can save large numbers of lives and do so at a high cost. Ever since the Reagan administration, American presidents have focused on producing 'net benefits' or benefits minus costs. If the net benefits are high, we have good reason to go forward, whatever our abstract misgivings about regulations (Sunstein 2013, p. 33).

As we can see, Sunstein perceives it as meaningless to be against regulations *per se*. The characteristic feature of his approach, however, is that it attempts to move us beyond the unhelpful notions of regulations, either as something inherently good, as the political left tends to believe, or as inherently evil, as the political right tends to believe. The problem, he thinks, is that the political debate easily gets stuck in less than helpful categories focusing either on "the dangers of 'more' government" and "the risk of socialism" or on "threats posed by 'the big polluters' and 'the banks'" (Sunstein 2013, p. 35). Sunstein is a pragmatist – as is President Obama (Kloppenberg 2011). He therefore thinks that the best way to get out of stalemates of this kind is by way of testing our beliefs. "What is needed to get out of this stalemate not only in the United States but all over the world, is smart, innovative strategies and tools, focused above all on evidence and on what works and what doesn't" (Sunstein 2013, p. 35).

This appeal to empirical testing and to "what works and what doesn't" is, as we have previously seen, the very trademark of pragmatism. What Sunstein is suggesting, therefore, is simply that both the political left, with their idea of regulation as something inherently good, and the political right, with their idea of regulation as something inherently bad, should subject their ideas to testing. What are the effects of a specific instance of regulating? Let's find out! And – not least – let us be willing to reconsider our viewpoints in light of what we find!

The basis for Sunstein's approach is found in the field of study which is commonly referred to as "behavioral economics". This field, which was founded by Amos Tversky and Daniel Kahneman, is characterized by the attempt to merge economic theory with the findings of empirical psychology. The assumption underlying the field as a whole is that mainstream economists are mistaken in identifying man as a rational agent. It may be helpful, already at this point, to remind ourselves of pragmatism's criticism of the assumption that we, in fact, are self-interested beings. The point, namely, is that they both draw the same conclusion from this finding. Both behavioral economics and pragmatism may be said to imply that *because* we are not purely rational – *because* our own principles and beliefs are not simply to be trusted – we ought to approach these principles and beliefs as fallible, and subject them to testing.

According to behavioral economists it is helpful to draw a distinction between two distinct modes of thinking, one intuitive and automatic, one reflexive and rational. These two modes may, with Daniel Kahneman (2011), be referred to as "System 1" and "System 2" respectively. According to Kahneman it is "System 1" that we make use of when we duck when an object is thrown at us, or smile when we see something cute, or complete a sentence such as "bread and ...". All of this is something that we do immediately, without consciously thinking. System 1 works fast, and is experienced as intuitive. It does not even involve what

we normally associate with thinking. This, on the other hand, is precisely what is characteristic of "System 2", which is what we apply when we decide to check whether an argument is valid, when we decide which road to take, or decide on an education (Kahneman 2011).

Both of these two systems are active as long as we are awake. "System 1" continuously generates suggestions to System 2. If System 2 accepts these suggestions, our impressions and intuitions are turned into beliefs, impulses are turned into actions. We do not always bother, however, to run our intuitions through System 2. For the most part, this too turns out well. Generally speaking, the impressions we receive are trustworthy and we can act on the basis of the desires and wishes that we have. But this is not always so. System 1 is good at dealing with well-known situations and hits the mark fairly well when it comes to making short-term predictions. In addition, its reactions are, for the most part at least, quick and to the point.

System 1 nevertheless makes systematic mistakes. It makes use of misleading rules of thumb, and "jumps" to conclusions which do not follow from the premises. It is heavily influenced by how different alternatives are presented to us, and sometimes even by information which is completely irrelevant to the decisions that we are about to make. What this implies is that how much we eat in fact will depend on how large the portions are which we are served. What we choose to buy is dependent on how the different alternatives are presented to us. We are more apt to choose a medical procedure which has 20 per cent change of succeeding than one that has an 80 per cent chance of failing. We believe that it is less likely that *we* will be the victims of an accident, even when it comes to something completely contingent, such as being struck by lightning. We evaluate the worth of objects differently, depending on whether it is something we ourselves own or whether it is merely something that we want. Smoking, to be sure, is the subject of a number of systematic errors of this kind. Even if we are aware of the dangers of smoking, we do not think that we will be affected by them. Or we believe that since so many people smoke, it cannot be as dangerous as the medical experts claim. Or that since we know someone who has smoked all her life without getting cancer, we are not likely to get it either. And so on.

These findings, to be sure, confirm the basic presuppositions of the pragmatists. Both Peirce and Dewey, we have seen, argue that any attempt to "secure" inquiry by ridding us of all unwarranted assumptions is doomed from the start. This is because, as they see it, inquiry takes place against a background of already established practices. It is this background that "sets" the problem. In our day-to-day life, therefore, we do not have ready answers as to why we do things in precisely the manner we do. It is just the way we do things. We sim-

ply rely on socially established procedures, habits etc. What the research of the behavioral economists brings forth, however, is that these procedures result in certain *systematic* mistakes. The point is thus not only that our established procedures *may* lead us into problems, as the pragmatists tended to remind us. The point rather, made by the behavioral economists, is that these established procedures *constantly run into problems* – and in highly systematic ways.

According to Sunstein, the empirical testing of principles and beliefs may thus have exactly the same function as that which System 2 has in regard to System 1. That is, it is only by way of empirical testing that we are able to keep these systematic mistakes at bay. Dewey would have agreed. As he sees it, the only alternative to testing our beliefs is "prejudice, the pressure of immediate circumstance, self-interest and class interest, traditional customs, institutions of accidental historical origin" (Dewey 1960, p. 265). To be against testing one's beliefs empirically is, both for Dewey and Sunstein, to be a dogmatist. There just isn't any way to make progress, they believe, without approaching our principles and beliefs as open to revision. This is the case in the empirical sciences, it is the case in politics, and it should also be the case in debates over moral issues.

V

Pragmatists, of course, tend to be rather optimistic about what may be achieved by way of their own approach. Dewey famously believed that by way of pragmatism, philosophers could once again concern themselves with the "problems of men" (i.e. with ameliorating the conditions under which we live), rather than merely with the problems of philosophers (i.e. with issues such as the "Why be moral?" question). Sunstein for his part, we have seen, believes that the pragmatic approach may serve as a "third way", thereby uniting the political left and the political right. This, it is safe to say, is overly optimistic. Anyone with the slightest knowledge of contemporary American politics, for instance, will be aware that it is not characterized by compromises.

However, this lack of success, which of course may be deemed somewhat ironic, given pragmatism's stress on "what works", need not be taken to indicate, however, that there are principled problems related to pragmatism as such. After all, for pragmatism to work, people must *be* pragmatists (rather than dogmatists). Pragmatism's lack of success may thus be explained, not with reference to problems in regard to *it*, but with reference, rather, to the dogmatism of its opponents.

This, however, will hardly be sufficient for convincing those who in fact oppose pragmatism. Has pragmatism of the Deweyan kind, they will ask, really succeeded in solving the presented dilemma? Or is its solution merely apparent? May it not *still* be objected, even *after* Dewey's revision, that pragmatism's emphasis on the good – or on "welfare" as we would say today – comes at a price, namely that of personal autonomy? Might it not still be insisted that which principles and beliefs I choose to adhere to must be up to me, not because leaving it up to me necessarily will have the best results, but precisely because *it will then have been my choice?* "If my choosing it is part of what makes my life plan good", Kwame Anthony Appiah states in *The Ethics of Identity*, "then imposing on me a plan of life – even one that is, in other respects, an enviable one – is depriving me of a certain kind of good. For a person of a liberal disposition, my life's shape is up to me, even if I make a life that is objectively less good than a life I could have made, provided that I have done my duty toward others" (Appiah 2005, p. 14).

We have previously seen that Christine Korsgaard argues that any successful answer to the question, "Why be moral?" "must show that sometimes doing the wrong thing is as bad or worse than death", and that "the only thing that could be as bad or worse than death is [...] not being ourselves anymore". Given this point of view, morality is intimately tied up with our own identity. So, it may be asked, doesn't pragmatism's focus on testing and effects simply miss out on what is truly important in regard to morality, namely that it is by way of being moral that we maintain our own self? May not that which the pragmatists themselves dismiss as "dogmatism", i. e. not to revise one's beliefs as a result of testing, be viewed as something *positive*, namely to stand by one's own principles and beliefs – and thus, in the last instance, by one's own self? Given Dewey's pragmatism, can anything possibly be "as bad or worse than death"? And if not, should this not be seen as a problem? Those of us who will answer this last question in the affirmative will, I suspect, still dismiss Dewey's pragmatism as an attempt to change the subject. *Rather* than focus on morality, which is an issue intimately related to our own identity and to our selves as agents, they will argue, Dewey is focusing on *prudence*.

It must be emphasized that the point is not that pragmatism cannot accommodate autonomy as such. The point, rather, is that it can only accommodate *weak* versions of it. For a pragmatist there *is* room for weak versions of autonomy. According to Dewey's revised approach, we have seen, notions such as 'duty', 'obligation' and 'right' must be analyzed independently of the 'good'; that is, they belong to a distinct sphere. Nevertheless, the question both can and ought to be raised whether that which we *should* do (because it is our obligation or duty) in effect results in something good. According to Sunstein, it is a

fact that we appreciate autonomy. As he sees it, however, this merely forms part of our total welfare. Autonomy may thus be viewed as one of the (many) factors that should be included in the cost-benefit analysis (Sunstein 2014, p. 124).

Against the backdrop of what has been said, the relative lack of success that the pragmatic approach has had in the practical-political field, and especially in ethics, may be taken to indicate that many of us are not satisfied with weak versions of autonomy. Those of us who are not will see autonomy, not merely as a part of welfare but as an end in itself and thus as decisive. They will argue, with Appiah, that the point of my life being up to me is not that it will result in something objectively better than if it were not, but that it is *up to me*. This is the *strong* version of autonomy.

The problem with the strong version of autonomy, from the point of view of Dewey's pragmatism, is that it turns out to be empty; it doesn't have any content. As we have seen, Dewey points out that it is our desires and needs that enable us to experience something as something. Any notion of the self which implies that we can step back from these – and this is precisely what is implied by the strong version of autonomy – must consequently be dismissed.

Defenders of the strong version of autonomy, however, may argue that this is a notion which must be upheld *nevertheless*. The reason for this is that it may be seen as forming a prerequisite for seeing what we do and say as actions. How, they will ask, is it possible to perceive that which I do as actions, as something which I am responsible for, unless these are said to have their "origin" or "source" precisely in me? On this basis, the strong notion of autonomy may be said to accommodate a certain perspective on ourselves and our doings – that my doings have their "origin" *in me* – which is necessary for us to presuppose, even in order to make sense of fields such as ethics, politics and jurisprudence. On this basis, defenders of strong autonomy may turn pragmatism's dismissal of the "Why be moral?" question on its head. Since pragmatism's empirical approach cannot account for the strong version of autonomy, it may now be argued, and since this notion forms a condition for the very possibility of perceiving us as acting beings, it isn't this *issue*, but *pragmatism's dismissal of it* that must be dismissed.

The unavoidable conclusion, therefore, is that pragmatists seem to be drawn right back into the dilemma which they believe themselves to have found the solution to. We thus still find ourselves in the position of having to choose between one view, advocated by Kant, Habermas, Appiah and others, according to which the notion of a distinct moral sphere must be upheld (since any attempt to dismiss it will have the result that we miss out on morality altogether), and another view, advocated by pragmatists of Dewey's kind, according to which the notion of a distinct or autonomous moral sphere must be dismissed (since any such no-

tion will be empty). Whether we believe that the "Why be moral?" question can be dismissed, therefore, will depend on which horn of the dilemma we choose. Both horns will leave us with serious objections.

Bibliography

Appiah, Kwame Anthony (2005): *The Ethics of Identity*. Princeton: Princeton University Press.

Dewey, John (1978): *Ethics 1908. The Middle Works of John Dewey 1899 – 1924. Volume 5*. Carbondale. Southern Illinois Press.

Dewey, John (1960): *The Quest for Certainty*. New York: Capricorn Books.

Dewey, John (2002): *Human Nature and Conduct*. New York: Prometheus Books.

Dewey, John (1985): *John Dewey: The Later Works, 1925 – 1953. Volume 7: 1932*. Carbondale: Southern Illinois Press.

Habermas, Jürgen (2005): "Norms and Values: On Hilary Putnam's Kantian Pragmatism". In Habermas, Jürgen: *Truth and Justification*. Edited and with translations by Barbara Fultner. Cambridge, Massachusetts: The MIT Press.

Hickman, Larry A. (2007): *Pragmatism as Post-Postmodernism*. New York: Fordham University Press.

Kahneman, Daniel (2011): *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kloppenberg, James T. (2011): *Reading Obama: Dreams, Hope and the American Political Tradition*. Princeton and Oxford: Princeton University Press.

Pappas, Gregory Fernando (2008): *John Dewey's Ethics: Democracy as Experience*. Bloomington and Indianapolis: Indiana University Press.

Peirce, Charles Sanders (1955): "The Fixation of Belief". In Buchler, J. (ed.): *Philosophical Writings of Peirce*. New York: Dover Publications.

Rorty, Richard (2007): "Kant vs. Dewey: the current situation of moral philosophy". In Rorty, Richard; *Philosophy as Cultural Politics. Philosophical Papers Volume 4*. Cambridge: Cambridge University Press.

Sunstein, Cass R. (2013): *Simpler: The Future of Government*. New York, London, Toronto, Sydney, New Dehli: Simon & Schuster.

Sunstein, Cass R. (2014): *Why Nudge? The Politics of Libertarian Paternalism*. New Haven & London: Yale University Press.

Twain, Mark (1986): *The Adventures of Huckleberry Finn*. Harmondsworth, Middlesex, England: Penguin Classics.

Wellmer, Albrecht (1998): "Truth, Contingency, and Modernity". In Wellmer, Albrecht: *Endgames*. Cambridge Massachusetts and London, England: The MIT Press.

Héctor Wittwer

Is the Overridingness of Moral Reasons a Semantic Fact?

At first view, the question, "Why be moral?" seems to allow for many intelligible answers. On closer consideration, though, there are only a handful of suitable responses to this question. As I have argued elsewhere (Wittwer 2010), all the different versions of these answers can be subsumed under only two or three types. According to the first kind of answer, we should always—or, at least in most cases—act morally because this serves our enlightened self-interest. This is what Plato had in mind when he tried to demonstrate in the *Republic* that being just is a necessary and sufficient condition for achieving happiness.[1] In the twentieth century, ethical contractualists such as David Gauthier and Gregory Kavka held a similar position: For enlightened rational egoists it is rationally prescribed to submit themselves to the constraints of morality in order to maximize the realization of their preferences: "To choose rationally, one must choose morally. [...] Morality, we shall argue, can be generated as a rational constraint from the non-moral premises of rational choice." (Gauthier 1986, p. 4)—Whereas the first answer to the question tries to reconcile the requirements of morality and of self-interest—this is what Kavka called the "reconciliation project" (see Kavka 1984)—the second kind of answer is based on the assumption that there is no need to show that moral actions are useful in any sort. On the contrary, the mere attempt to demonstrate that the good and the useful are identical proves that the concept of morality has been misunderstood. As F. H. Bradley has argued in his seminal paper "Why Should I Be Moral?", the good is an end in itself, that is, something which we ought to do for its own sake whichever consequences this may have with regard to our self-interest. Therefore, the question "Why should I be moral?", according to Bradley, "seems strange. For morality (and she, too is reason) teaches us that, if we look on her only as a good for something else, we never in that case have seen her at all. She says that she is an end to be desired for her own sake, and not as a means to something beyond." (Bradley 1990 [1876], p. 53) We always ought to act morally because morality is the only *end in itself*. Some authors have grounded this claim on the assumption that morality is the only thing which has an *absolute inherent*

1 "Then the just is happy, and the unjust miserable? – So be it." (*Republic*, 354a [translation by Benjamin Jowett], 1892).

worth. For example, in a well-known passage from the *Groundwork of the Metaphysics of Morals* Immanuel Kant asserts

> [...] nothing has any worth other than that which the [moral] law determines for it. But precisely because of this, the legislation that determines all worth must itself have a dignity, i. e. unconditional, incomparable worth, for which the word *respect* alone makes a befitting expression of the estimation a rational being is to give of it. *Autonomy* is thus the ground of the dignity of a human and of every rational nature. (Kant 2012 [1785], pp. 47–48 [AA IV: 436])

Thus, there are at least two prominent answers to the question "Why be moral?" According to the first, we should act morally because this is in our enlightened self-interest. According to the second, however, we should be moral because morality is an end in itself or the only thing which has absolute inherent worth.

It might be that the alternative between these two types of answers is exhaustive. This depends on whether the answer to which this paper is dedicated can be subsumed under the second one. Although this is an important theoretical issue, I cannot explore it here. Instead, I will present the remaining answer without trying to decide whether or not it is only a version of the second one.

The response I am going to present and critically evaluate says, firstly, that we ought to be moral because moral reasons are overriding and, secondly, that the overridingness of moral reasons is a *conceptual feature of morality*. In other words, it is part of the common understanding of the concept of morality that there can never be any sufficient reason for not doing what one knows to be morally prescribed. Everyone who has learnt how to properly apply terms as "moral", "morality", "ought" or "the good" has thereby aquired the belief that we always ought to do what is morally required whatever other reasons may speak against doing it. The overridingness of moral reasons is said to be a part of the language-game of morality and, hence, a *semantic fact* (see Hoffmann 2013, p. 595). In other words, "overridingness is genuinely part of our pre-theoretical conception of, or at least aspirations for, morality" (Stroud 1988, p. 176). The assumption that moral reasons are always normatively overriding is said to be an "analytic truth" (Hoffmann 2013, p. 595).

This assertion has an important implication. If it is true, then it is impossible to reasonably ask "Why be moral?" because the question "Why should I do what is morally required?" can be translated into the question "Why should I regard the only kind of reasons which are, by definition, overriding as overriding?" This would, indeed, be a pointless question. By asking it, the speaker would give proof of his deficient understanding of the word "morality". Therefore, the only adequate reaction to his request would consist in explaining to him the very meaning of words such as "ought", "morality" or "the good". Once

he has fully understood the meaning of these words, he can no longer reasonably ask question, "Why be moral?" In a well-known passage F. H. Bradley expressed this line of thought in the following way: "Has the question, Why should I be moral? no sense then, and is no positive answer possible? No, the question has no sense at all; it is simply unmeaning, unless it is equivalent to, *Is* morality an end in itself, and, if so, how and what way is it an end?" (Bradley 1990 [1876], p. 59)

The thesis that overridingness is a conceptual trait of morality and thus a semantic fact was most famously held by *R. M. Hare* in the second half of the last century. Many philosophers have followed him. According to Hare, besides *universalizability* and *prescriptivity*, overridingness is a *logical property of moral judgments* (see Hare 1981, p. 24 and pp. 50 – 61). The strength of Hare's response to the sceptical challenge that underlies the question "Why be moral?" is that it is completely independent of empirical issues. Whatever the social world we live in might be like, moral reasons will always have priority over all other normative reasons for actions. We do not need any empirical information about human beings and human societies in order to decide whether or not moral reasons are overriding. Instead, our knowledge of the overridingness of moral reasons is already part of our linguistic competence.

Before presenting my objections against this view, it seems appropriate to mention five of its aspects in order to avoid misunderstandings of the position held by Hare and his followers.

(i) First of all, OT—as I will call the overridingness thesis from now on—does not apply to single moral reasons taken in isolation from all other relevant features of a given situation but to overall moral judgments about what the agent ought to do. Therefore, the moral reasons which are said to have priority over all other kinds of normative reasons have already taken into account all competing *prima facie* duties, exceptions and excuses. Thus, OT says that if doing X is, *all things considered*, morally required, then there can be no sufficient reason for refraining from doing X (see Stroud 1988, pp.172 – 173). If we want to construe cases in which there is a conflict between self-interest and morality, we have to take this aspect of OT into account because, if we do not, we will include situations in which there is no real incompatibility between prudential and moral reasons. Let me mention just one example that has been discussed and misinterpreted in the literature on overridingness: after World War II many Germans who were suffering from hunger used to steal potatoes from trains or stores in order to feed themselves and to survive. It seems very plausible to me that these people were morally excused because they found themselves in a situation in which they were faced with the choice between starving to death and violating the moral rule which forbids stealing. Thus, the case is not an example of a conflict

between self-interest and morality. If we assume that a violation of the moral rule which forbids stealing is excused if stealing is necessary in order to survive then the potato thieves acted in a morally permissible manner. Therefore, the example does not prove that reasons of self-interest can have normative priority over moral reasons. Rather, it illustrates that it can be morally permissible, or at least morally excusable, to violate a moral rule if this is necessary in order to avoid an evil which is clearly and significantly worse than the one which is caused by breaking the rule. This, however, has rarely been denied either by proponents or by critics of OT.

(ii) Secondly, we have to distinguish between two kinds of overridingness: *actual* and *normative* overridingness. Practical reasons are *actually* overriding if an agent or a group of agents always gives them priority over all other kinds of reasons. Practical reasons are *normatively* overriding if all moral agents always ought to give them priority over all other relevant considerations. It is obvious that OT does not refer to the actual overridingness of moral reasons. As we all know there is no doubt about the fact that, often, we and other people do not treat moral reasons as overriding. However, as OT refers only to the normative priority of morality, this fact does not speak against it. OT cannot be falsified by empirical evidence about real decisions of moral agents.

(iii) The third aspect is based on another conceptual distinction. This time we have to distinguish between two kinds of normative priority or overridingness. If a reason takes normative precedence over another one *only with regard to a certain criterion or a certain domain of decisions and actions*, then I will speak of *relative* priority. For example, it cannot be denied that, in case of a conflict between my desires and the requirements of morality, with regard to my self-interest my prudential reasons have relative priority over the moral reasons. Or, to give another example, if the law in a certain country forbids something which many of its citizens regard as morally mandatory, then both the law and the respective moral rule can be normatively overriding depending on which criterion one uses for solving the conflict. With reference to the state and its judicial system, the law has priority; but with regard to the moral integrity of the citizens the moral rule should be overriding. Now, it should be clear that according to OT moral reasons are not only relatively overriding but that they always have *absolute* priority over all other kinds of reasons.[2] Of course, it is true that, judged from the moral point of view, moral reasons should take precedence over prudential reasons. But, unless it is demonstrated that one ought to solve the conflict from

2 Although she does not use the concept of absolute overridingness, this aspect is stressed by Sarah Stroud (1988, p. 175 – 176).

the moral point of view, this only shows that moral reasons are relatively overriding. Now, this is obviously true with regard to all types of normative reasons. Hence, relative priority is not a distinctive feature of moral reasons.

(iv) The fourth aspect of OT is closely linked to the former one. It is assumed that absolute normative priority is *a distinctive feature of moral reasons*. This assumption can be split into to sub-claims: (a) If a normative reason for action or the corresponding norm is not absolutely overriding, it is not a moral reason or norm. (b) If a normative reason for action or the corresponding norm has absolute priority, then we are dealing with a moral reason or a moral norm. From these alleged features of morality some authors have formally correctly inferred that moral reasons and norms can be *identified* by reference to absolute overridingness (see Kuhlmann 1985, p. 187).

(v) Finally, I would like to point to an implication of OT which is of the utmost significance with regard to the question "Why be moral?": if the absolute overridingness of moral reasons is part of the meaning of terms like "morality" or "ought" then it seems that there can never be a sufficiently good reason for not doing what is morally required. This is tantamount to saying that, in one sense or another, it can never be rational not to do what is morally prescribed. This claim can be understood in two ways. According to its *strong* version not giving priority to reasons which are normatively overriding is *self-contradictory*. As striving for consistency is the most fundamental requirement of rationality, it is always and necessarily irrational to act against a correct overall moral judgment. According to the *weak* interpretation, however, acting against the precepts of morality, although it is usually rational, can never be *fully* rational. In this case, it is assumed that persons who break moral rules generally do so for prudential, political, religious or other non-moral reasons. As acting for reasons cannot be wholly irrational violations of the moral rules are not wholly irrational. On the other hand, the person who infringes a moral rule does not act on the *best* reasons available to her. Therefore, she makes only a *deficient* use of practical reason. Hence, her decision and her action are *not fully* rational. Now, if we link these assertions to another one which seems to me to be undeniable, namely that we always should act as rationally as possible, then we can give a very strong answer to the question "Why be moral?": we should always do what is morally required because, unless we do so, we will act irrationally or, at least, not fully rationally.

Having clarified what the claim that the overridingness of moral reasons is a semantic fact implies, I will now proceed to an evaluation of this assumption. First of all, it is evident that there is no consensus about the meaning of the word "morality" among all competent speakers of languages such as English or German. The very proof of this is the fact that philosophers have been debat-

ing for a long time about whether or not overridingness is a distinctive feature of moral reasons. Hence, there is disagreement about the meaning of the concept of morality. Now, it is not unusual that there is a partial disagreement about the meaning of certain words within a linguistic community. Such differences in opinion can be resolved if at least one of the following two conditions is met. If the disputed term applies to a *natural kind* then there is an objective criterion for the correct definition of the term. In this case, the direction of fit goes from the meaning of the word to the natural kind it refers to. If the former does not correspond to the latter, it has to be modified. For example, the adequate definition of the concept of a tiger depends on what tigers are like. The second option for resolving a semantical disagreement presupposes a certain level of division of labor. If there is a group of specialists, for instance scientists, who due to their professional training and activity have some special knowledge about the disputed concept at their disposal which all other speakers lack, then it is up to these specialists to determine the correct meaning of the word. If, for example, we as laymen do not agree about the proper application of words like "polis" or "agora", we have to address ourselves to historians or classical scholars in order to resolve our dispute.

Unfortunately, none of these options is available with reference to the word "morality". Morality is neither a natural kind, nor is there a group of specialists endowed with a knowledge about morality to which other people have no access. Let me briefly explain these two claims. Firstly, morality is a social order or a social practice, hence something which cannot exist independently of how human beings conceive of it. If there were no human beings or creatures sufficiently like us, there would be no social institution like morality. Therefore, there is no objective criterion for determining the adequate meaning of the word "morality" which is prior to our convictions about what morality is. Secondly, considered as moral agents, with regard to the meaning of moral concepts all normally developed human adults are on an equal footing. No one has a priviliged access to the meaning of moral terms. Hence, there is no group of specialists to which we could turn in order to figure out whether or not absolute normative overridingness is a conceptual and, thus, a necessary feature of morality.

Given these premises, what can we do in order to decide the issue? All we can do is listen as carefully as possible to ordinary people talking about moral issues and, in particular, about conflicts between self-interest and morality. If we do this, we will certainly find out that absolute normative priority is not part of the common meaning of words like "morality" or "ought". Although it is true that people often ascribe normative priority to moral reasons in their judgments, it is not at all inconsistent to accept the limited authority of morality and to deny that it is absolutely overriding. This fact is easily overlooked because

many philosophers do not distinguish between two kinds of utterances: *sincere* utterances on the one hand and *insincere* utterances made by speakers who do not give absolute priority to moral reasons but publicly pay lip-service to the alleged authority of morality on the other hand. Now, it is obvious that under usual circumstances nobody who is willing to treat prudential reasons as overriding with reference to moral reasons will *publicly* declare that he does not accept the claim that morality has absolute normative priority. The reason is evident: Publicly declaring that one is willing to give priority to self-interest over moral requirements would simply be foolish. If, for example, I want to benefit from the social practice of giving and keeping promises I should not declare that although I profit from the general practice of giving promises, I am willing to break my promises if that will serve my interests. By doing this I would lose the confidence of those who want to rely on promises. They would no longer regard me as a potential co-operator. It is in my interest that I be considered as someone who will always keep his promises even if this involves considerable losses or sacrifices.

From this example, which is representative for every moral practice or institution, an important lesson can be drawn: the fact that usually nobody *publicly* denies that moral reasons are normatively overriding does not prove that overridingness is a conceptual feature of morality. It only shows that generally everybody has accepted that it is imprudent to declare in public that other kinds of reasons can have priority over moral demands. As Plato had Glaucon say over two thousand years ago: "the highest reach of injustice is, to be deemed just when you are not".[3]

Things look quite different if we turn to the *private* use of language. Behind closed doors people often explicitly state that, in case of a conflict between morality and self-interest, one should give priority to prudential reasons. This does not only hold for criminals who, by the way, usually know very well that what they do is morally prohibited. Sometimes, ordinary people like you and me also judge that somebody should not do what he is morally obligated to do because it is not in his interest. Judgments of this kind are not contradictory; at least, they do not look like they involve a contradiction at first sight. Imagine the following case: A relatively young man who works in a private company has witnessed that a female colleague of his whom he knows quite well has been dismissed under a pretext because she had not reacted positively to her boss's sexual advances. Now the young man's wife tells him that although it is his moral duty to protest against this unjust decision he should keep quiet

3 Plato, *Republic* 361a (translation by Benjamin Jowett 1892).

and not mention the incident because this would endanger the promotion he has been waiting for. Compare the sentence "Although it is your moral duty you should not do it because it would be imprudent" to utterances which are obviously contradictory, such as "This circle has two angles" or "All immigrants are lazy but my Turkish friend who owns the greengrocery at the corner is a hard-working man".[4] Unlike the two latter sentences, the former seems to be in no way inconsistent. But, if this is true, absolute normative overridingness cannot be a conceptual feature of morality. If absolute normative priority were indeed a conceptual feature of overall moral reasons, the young man's wife's utterance could be translated into "You should not do what you ought to do" or "Although there can never be a sufficient reason not to do what you are morally obligated to do, you should not do it". But, it seems to me that what the young man's wife said cannot be accurately translated in this way.

Let me briefly summarise my first objection to the claim that the overridingness of moral reasons is a semantic fact. Up to this point, I have argued that absolute normative priority is not an inseparable part of the common understanding of the concept of morality. I have based my objection on the conceptual distinction between public and private utterances and on the observation that sentences in which the normative overridingness of moral reasons is explicitly denied do not seem to be contradictory. Although I am sure that everything I have said so far is correct, I fear that some readers might not be convinced yet because, as we all know too well, philosophers tend to be extremely sceptical. Therefore, I am going to present a second objection which is completely independent of the first one.

The point of the first objection was that the mere fact that somebody denies the absolute normative priority of morality and moral reasons does not prove that this person has a deficient understanding of terms like "morality", "ought" or "the good". The feature of normative overridingness is not an integral part of the meaning of "morality". Now, let me assume from now on for the sake of argument that all competent speakers would agree that overridingness is a necessary feature of overall moral reasons for actions, hence that everybody would consider moral reasons as overriding. Would this semantic consensus prove that moral reasons are really absolutely overriding?—I do not think so. In what follows I will try to explain why the overridingness of morality cannot be a semantic fact.

In order to explain my thesis I have to make a short digression into the philosophy of language. Every accepted meaning of a concept or of the words which

4 This example is intended to illustrate the fact that moral beliefs can be self-contradictory.

denote that concept in different languages can be considered as a definition or, at least, as a part of the definition of that concept. If, for example, all competent speakers agree that the concept of air refers to an invisible mixture of gases then invisibility and being gaseous are parts of the universally accepted definition of the concept of air. Now, at least some definitions can be true or false insofar as they can be adequate or inadequate with regard to the objects to which they refer. If air were not a mixture of gases but a liquid then the universally accepted meaning of the word "air" would be inadequate. There are numerous examples of meanings which were universally accepted and nevertheless inadequate. For many centuries "being a fish" was an integral part of the word "whale". Therefore, whales were called "Walfische"—"whale-fish"— in German. Hence, the feature of being a fish was an integral part of the universally accepted definition of whales. Of course, nowadays we know that whales are not fish, but mammals. Thus, the word "whale" was used for a long time implying an inadequate meaning.

The upshot of this line of thought is that one must not without further ado infer from the universally accepted meaning of a word claims about the nature of the object which is denoted by that word. Accepted meanings are definitions or parts of definitions and, thus, can be adaquate or inadequate, true or false. If this claim also holds for the concepts of morality and of moral reasons then the alleged fact that all speakers agree that overridingness is an integral part of their meaning is not a sufficient reason for concluding that absolute normative priority is indeed a necessary feature of morality itself.

However, at this point we are faced with a problem which arises because of the social character of morality. As I pointed out earlier, morality is not a natural kind, but a social practice which cannot exist independently of what people take it to be. Therefore, proponents of the overridingness thesis could argue in the following way: Morality is what competent speakers take it to be. What the nature of morality is cannot be decided independently of what we understand by the concept of morality. Now, all competent speakers agree that absolute normative priority is a conceptual feature of morality. Hence, absolute overridingness is, indeed, a necessary trait of morality. They could illustrate this argument by pointing to an important difference between the concept of morality and the concept of action. Outside departments of philosophy, almost everybody takes for granted that human decisions are not completely determined by natural causes. It is generally held that, unless somebody is forced to act in a certain way, he or she can always freely choose between at least two options, namely performing a certain action or refraining from doing it. Having the choice between alternate possibilities is, thus, an integral part of the universally accepted meaning of the pre-theoretical concepts of decision and action. However, even if this is

true it does not follow that we do have a free will because the meaning that we ascribe to the words "decision" and "action" could be inadequate. In order that our will be free, an external condition must be met. Only if the events in the world are not completely determined by natural laws—that is, only if there are alternate possibilities—can we freely decide to act one way or another. Whether or not we are free does not only depend on what we understand by the concepts of "decision" and "action", but also *on the way the world is*. But, the proponent of OT might continue, if we talk about the nature of morality we are in a completely different situation. Morality is what human beings think it is. We need not know the natural constitution of the world in order to tell whether a certain trait is a necessary feature of morality. As all competent speakers agree that overridingness is a conceptual feature of morality, moral reasons are, indeed, overriding. The overridingness of morality is simply a semantic fact.

This argument in favour of OT may sound compelling. But it is based on an all too simple premise. Our moral concepts and our concept of morality itself are not isolated elements of the languages we speak. The beliefs that are incorporated in them are part of a whole network of beliefs about actions, their possible justification and about practical rationality. Therefore, proponents of OT have to demonstrate that their claim fits into this web of beliefs. There is at least one belief about the justification of actions which is undoubtedly true and which has to be reconciled with OT because there is a certain tension between this belief and the alleged priority of morality: moral reasons are not the only kind of normative reasons for action. There is at least one other type: prudential reasons. For the purposes of my argument, the question of how many types of normative reasons for actions there are can be left open. Suffice it to say that moral reasons are not the only kind of reasons which can justify an action. Therefore, conflicts between different kinds of normative reasons can arise. OT is true if and only if it is rationally mandatory to act on moral reasons whenever there is a conflict between moral reasons and reasons for action of a different kind. (In fact, OT is only understandable if one assumes that there are at least two types of normative reasons for action. If there were not, there would be nothing with regard to which moral reasons could have priority.)

Let me briefly explain this assumption. As far as I know, nobody has ever claimed that moral justification is the only possible kind of justification. Usually, we do not have to give moral reasons for what we do. Most of the decisions that we have to take day by day have nothing to do with moral problems. Nevertheless, we often must deliberate, decide, and sometimes also justify our decisions. In many cases, we do this with recourse to the criterion of our enlightened self-interest. The ability to make choices with recourse to this criterion and to act according to these choices is a type of practical rationality which shall be called

"prudential rationality". The demands of prudential rationality are no less normative than those of moral rationality. It cannot be denied that conflicts may arise between the requirements of prudence and those of morality. In order to demonstrate that moral reasons are normatively overriding, it has to be shown that persons who have to make choices between moral and prudential reasons always ought to act on the moral reasons. It is plain that this task cannot be fulfilled by simply asserting that normative overridingness is part of the universally accepted meaning of the concept of morality, for this would be an obvious *petitio principii*. Instead of justifying the linguistic usage, one has simply presupposed its correctness. What is needed if one wants to avoid this *petitio* is a justification which is beyond the scope of the mere meaning of certain concepts, that is a *non-linguistic justification*. This is tantamount to say that even if normative overridingness is a necessary feature of moral reasons, this normative priority of morality cannot be a mere semantic fact.

Can there be a plausible non-linguistic justification of the alleged overridingness of morality? I suggest that this is impossible because all attempts to achieve this goal lead into the following *dilemma*. Either OT is restricted to the true but trivial claim that moral reasons are relatively overriding with regard to the moral point of view, or it refers to the absolute priority of morality. In the latter case the argument is doomed to fail because, as I hope to show in what follows, nothing can ever have absolute priority over something else.

Of course, these strong claims need an explanation. Let us consider the first horn of the dilemma. OT can be so modified that it only applies to the relative moral priority of moral reasons. In this case, it states correctly that, judged from the moral point of view, moral reasons for action should always take precedence over all other kinds of action. This assumption is certainly true; but it is also *trivial*. What is worse is that it does not provide us with a satisfying answer to the question, "Why be moral?" unless it is shown that one should resolve conflicts between competing reasons for action with recourse to the criterion of morality. But, of course, this is not implied in the weak version of OT. This version only states that *if* the conflict is judged from the moral perspective then moral reasons are overriding. But it cannot tell us why we should always pass judgments from the moral point of view. Hence, the weak version of OT is unsuitable for giving a satisfying answer to the question, "Why be moral?"

The strong version of OT faces another problem. According to it, moral reasons are not only overriding with recourse to a certain criterion. Instead, they are said to have absolute priority. What this means depends on how the ambiguous expression "absolute priority" is understood. That A has normative priority over B can mean that A is normatively overriding with regard to *all pertinent criteria*. It is evident that moral reasons cannot have normative priority in this sense be-

cause they are not overriding with recourse to at least one criterion, namley the criterion of prudence. According to the second possible interpretation of the strong version of OT there is *no criterion at all* with regard to which moral reasons are overriding. Rather, they are overriding *simpliciter* or *tout court*. Recently, Owen McLeod has defended this claim and added that there is a kind of ought which corresponds to the alleged overridingness *tout court:* he calls it "just plain ought" (see McLeod 2001). It should be clear that this is only an *ad hoc* solution which barely covers the argumentative weakness of the proposal. "Takes priority" is a three-place predicate. Something takes priority over something else with regard to a certain criterion. A takes priority over B with regard to C. This is reflected in the fact that it is always appropriate to respond to the claim that A takes precedence over B by asking: with regard to what criterion? The talk of an overridingness *tout court* and of a corresponding "just plain ought" is nothing more than a helpless attempt to arbitrarily cut off the philosophical discussion.

The second horn of the dilemma could only be avoided by demonstrating that there is a normative criterion for solving conflicts between morality and prudence which always tells us to act for moral reasons. This general point of view cannot be identical either with the moral point of view or the perspective of prudence. Why not? These two criteria are not only situated on the same level. They also refer to the same kind of objects, namely single decisions or actions. Finally, neither of them is neutral with reference to their conflict. Hence, the standpoint of morality and the standpoint of prudence are on equal footing. None of them is suitable for serving as an overall criterion for resolving conflicts between themselves. Therefore, OT could only be saved if there were a *rational meta-criterion* for resolving conflicts between different kinds of normative reasons for action.

But we do not possess such a content-neutral, rational meta-criterion. Hence, in case of a conflict between morality and prudence neither the moral nor the prudential action is rationally prescribed. In such a case practical rationality cannot tell us for what kind of reasons we ought to act. As there is no normative meta-criterion for choosing between them, both are rationally allowed. It is as rational to act morally as it is to act prudentially. We cannot solve the conflict between the two kinds of reason from a general point of view simply because this general normative perspective does not exist.[5]

It follows from the foregoing that the claim that moral reasons for action have non-moral priority over other kinds of reasons cannot be justified. The over-

5 A more detailed argument for these claims is to be found in Copp 1997, esp. pp. 100 – 105.

ridingness thesis is either true but trivial or false, and its true version cannot provide us with a satisfying answer to the question "Why be moral?".

At the close of my paper, I would like to discuss briefly an objection to what I have said. At first sight, it might seem that my position is unable to explain a fact which cannot be denied. People do actually pass overall judgments on conflicts between competing kinds of reason. They choose between acting for moral or for prudential reasons. It might be claimed that the overridingness thesis offers the best or even the only explication of this fact. This is what Sarah Stroud seems to have in mind when she invites us to

> consider the fact that some of us actually *take* moral requirements to be overriding: we treat them as defeating other claims. If morality is indeed overriding, then there is no diffculty in understanding this practice: such agents are simply responsive to the true weight of practical reasons. But if in fact morality is not overriding, a commitment to honoring its demands seems rationally unmotivated. (Stroud 1988, p. 176)

This quote contains at least two different assumptions. First, Stroud claims that OT provides us with a plausible explanation of the fact that people do resolve conflicts between prudence and morality in favour of the latter. Secondly, she assumes that if morality were not overriding we could have no rational reason for resolving those conflicts. I readily admit the first part. But it must be added that OT is not the only plausible explanation of the fact that people can make choices between self-interest and moral demands. The situation is different with regard to the second claim which I take to be false. Why is this so? First of all, Stroud, like other authors, does not sufficiently distinguish between the two kinds of overridingness mentioned earlier: actual and normative overridingness. Denying that morality is normatively overriding does not imply that moral reasons cannot be actually overriding.

Of course, it cannot be denied that people actually often give priority to moral reasons. But, in order to explain this fact we do not have to assume that they decide to act for moral reasons because they are responsive to the "true weight" of moral reasons. They can make the same decision for different reasons, for example because they fear external or internal sanctions. In the end it all depends on how much significance the individual agent attributes to morality and to his own interests. Since in the case of a conflict between self-interest and moral demands both actions are rationally allowed, the two different actions can be motivated by rational reasons for action both if agents choose to act for moral reasons and if they decide to act for prudential reasons. Finally, their preferences will decide. But whatever they do, no normative conclusion can be drawn from their actual choice.

To sum up my argument: The alleged normative overridingness of moral reason is not a semantic fact. It is no fact at all.

Bibliography

Bradley, F. H. (1990 [1876]): "Why Should I Be Moral?" In: *Ethical Studies*. Reprint. Bristol: Thoemmes Antiquarian Books, pp. 53 – 77.

Copp, David (1997): "The Ring of Gyges: Overridingness and the Unity of Reason". In: *Social Philosophy & Policy* 14. No.1, pp.86 – 106.

Gauthier, David (1986): *Morals by Agreement*. Oxford: Clarendon Press.

Hare, R. M. (1981): *Moral Thinking: Its Levels, Method and Point*. Oxford: Clarendon Press.

Hoffmann, Martin (2013): "Thomas Buddenbrook und der Vorrang der Moral". In: Miguel Hoeltje/Thomas Spitzley/Wolfgang Spohn (Eds.): *Sektionsbeiträge des achten internationalen Kongresses der Gesellschaft für Analytische Philosophie e. V.* Duisburg-Essen: Universität Duisburg-Essen (online publication), pp. 593 – 606.

Kant, Immanuel (2012 [1785]): *Groundwork of the Metaphysics of Morals*. Ed. by Mary Gregor and J. Timmermann. Cambridge/New York: Cambridge University Press.

Kavka, Gregory (1984): "The Reconciliation Project". In: David Copp/David Zimmerman (Eds.), *Morality, Reason and Truth: New Essays on the Foundations of Ethics*. Totowa, NJ: Rowman & Allanheld, pp. 297 – 319.

Kuhlmann, Wolfgang (1985): *Reflexive Letztbegründung. Untersuchungen zur Transzendentalpragmatik*. Freiburg/München: Alber.

McLeod, Owen (2001): "Just Plain 'Ought'". In: *The Journal of Ethics* 5, pp. 269 – 291.

Stroud, Sarah (1988): "Moral Overridingness and Moral Theory". In: *Pacific Philosophical Quarterly* 79, pp. 170 – 189.

Wittwer, Héctor (2010): *Ist es vernünftig, moralisch zu handeln?* Berlin/New York: de Gruyter.

Alan Thomas

Williams on Integrity, Ground Projects and Reasons to Be Moral

This paper addresses the question, "Why be moral?" from the standpoint of Bernard Williams's moral and political philosophy. Schematically, Williams's answer to this question was that ethical reasons, like all an agent's practical reasons, must be grounded in those fundamental projects with which an agent is most deeply identified. As a corollary of this most general claim, it follows that to act from such ground projects is to act *from integrity.*[1] One of Williams's concerns, in addressing this question, was whether or not a priori arguments can be given that the contents of anyone's ground projects are shaped by rationally inescapable principles of morality. (For example, principles of the kind defended by Thomas Nagel in *The Possibility of Altruism* [Nagel 1970].) His answer was negative: there is no sound transcendental argument to the effect that the requirements of morality are constitutive of rational agency as such. That critique is based on Williams's conception of practical reasons as all "internal" in a sense I will explain.

A further aim of this paper is to extend the consideration of Williams's answer to our orienting question from its origin in a conception of practical reasons in general – as always "internal" – to his later political psychology. This late discussion presents his considered verdict on the nature of ground projects [Williams 2002, chapter 9]. It complements his earlier account of internal reasons and integrity and addresses a major line of concern about his whole argument strategy. This concern is one that must be addressed satisfactorily if his treatment of these themes is to be plausible.

The concern is this: for Williams's critics, the key bridging concepts that take us from ethics to politics, those of a ground project and of an agent's integrity, fall outside the scope of the moral completely. That is why, for them, his argument is puzzlingly weak. Williams himself added that whatever resonance his account of an agent's integrity was supposed to have, it was not that of a principle

1 Williams, as we shall see, connects this idea of integrity to that of identification with those ground projects from which an agent's reasons stem: "[an agent] is identified with his actions as flowing from his projects and attitudes .... [*H*]*is* actions and *his* decisions have to be seen as the actions and decisions which flow from the projects and attitudes with which he is most closely identified". [Williams, 1973, p. 100, pp. 116 – 117]

being refuted by a counterexample.[2] Yet his discussion of integrity produced a plethora of such counterexamples: of principled slaveholders and Nazis who exhibit the putative virtue of integrity. If such people can act from integrity, then the connection between integrity and acting in the light of moral reasons seems wholly contingent and external. For Williams's critics, absent a prior commitment to the shaping of the agent's deliberative field – that which she is prepared morally to so much as countenance – by a commitment to moral principle, appeal to an agent's ground projects arrives too late. Morality is not something, as Nagel put it, that an agent can "beg off": its source can be known a priori and it is a constraint on the projects of all agents.

Interpreted in this way, Williams's argument is shoehorned into the contrast between altruism and egoism. With this context in place, ground projects play the role of putatively ethical commitments whose fulfillment may play a role in determining the agent's *own* conception of the good, but whose moral credentials are yet to be determined because they may turn out to be egoistic or immoral. Vicious actions may fulfill the ground projects of the principled Nazi and thus be "good for" her, but this is no help in explaining what it is to be moral. This one way trading of intuitions gives rise to a persistent sense that the two sides in this debate are talking past each other – did Williams really overlook such obvious counterexamples to his thesis?[3] Williams has ignored the fact that practical agency is the ultimate "ground project" and the moral principles that constrain it are inescapable for all agents.[4]

In this paper, I extend consideration of Williams's arguments in ethics to the arguments of *Truth and Truthfulness* because only there does Williams connect the idea of an agent's ground project to the modern politics of identity [Williams 2002]. Examining this argument shows three things: first, that in spite of the criticisms that Williams's proposal received, it was one to which he remained committed. Secondly, this treatment of ground projects in the broader context

2 "I can hardly deny that I produced an integrity objection in the sense that I made an objection and it centred on the notion of integrity .... The objection did not, however, take the form of my trying to disprove a theory by counter-example, as much of the discussion has assumed. If the stories of George and Jim have a resonance, it is not the sound of a principle being dented by an intuition." [Williams, 1995a, p. 211]

3 The critic continues: does not William himself note the importance of the pre-commitments to the outcome of ethical deliberation involved in what the agent is prepared to countenance? Equally, does he not also stress the category of that which the agent finds unthinkable? If the very same emphasis on the pre-structuring of the agent's deliberative field appears in Williams's work and his critics objections to it, then once again this discussion seems to have set off on the wrong foot.

4 This is the defining thesis of meta-ethical constitutivism [Nagel, 1970; Korsgaard, 2009].

of what it is to be truthful to oneself – and the political ramifications of that ideal – re-iterates his view that to act from integrity is not to act on an executive minor virtue that can be assessed independently of any assessment of an agent's ends in action. (A good example of the latter is the admirable trait of cleverness [Wallace 2001].) So a slaveholder and a Nazi cannot act from integrity given the badness of their ends in action (while they could pursue those ends while exhibiting cleverness).

Thirdly, and most importantly, this late discussion reinforces the point that to act from integrity is to act in the light of a ground project that calls for *acknowledgement* on the part of others. Acknowledgement has two dimensions: the first is that a person is harmed if the ground project with which he or she most deeply identifies fails to be acknowledged by others. The second is that such a project must be, first-personally, as much found as made. As a stable commitment to an identity there is a role to be played in the explanation of such projects by voluntary choice. For example, such a choice enters into whether or not a person structures their life around such identifications as a "homemaker" [Thomas, 2009a]. However, the element of commitment to an identity implies that such identifications cannot be adopted or abandoned *merely* at will. This final point reinforces my central claim: ground projects, as Williams conceived of them, presuppose an ideal of mutual recognition foreshadowed by his earlier discussion of the "proleptic mechanism" involved in blame. In the background to the entire discussion, then, is a conception of ethical community in which reasons are received and given.[5]

It might seem important, in characterizing such a community, to explain at length Williams's distinction between the ethical and the moral. Unfortunately, for reasons of scope I cannot give this distinction the full discussion it merits. I will note only that Williams's critique of the Morality System both informs, and is informed by, his conception of practical reasons as internal [Skorupski, 2007]. I will examine that connection in the opening section of this paper. I will, for the most part, treat the question of why one ought to be moral as why one ought to

5 Arthur Ripstein notes, of this conception of a modern society, that "the very possibility of persons conceiving of themselves as free and equal, and having the appropriate concepts so to regard themselves is *a historical achievement*" [Ripstein, 2010, p. 683, emphasis added]. For a discussion of this point in connection with Williams's late turn to political philosophy see Thomas [2015]. The point itself, and Williams's uneasy relationship to it, is a reminder that Williams's late work was completed under the shadow of chronic ill health, that it is sometimes not all of a piece, and that it is hard to reconcile the republican strands in his political philosophy with his foundational account of human rights in the context of his political realism.

be committed to the ethical in the least demanding sense that Williams characterized as:

> the capacity shown, in some form or other, by humans in all cultures to live under rules and values and to shape their behavior in some degree to social expectations, in ways that are not under surveillance and not directly controlled by threats and rewards. [Williams 2002, p. 24]

One has a reason to be moral in this sense if doing so stems from one's fundamental ground projects. To act in the light of such projects is to act from integrity; this is not, however, to be understood as an exercise in mere self-assertion as reasonable ground projects demand acknowledgement from others.

1 Williams's Psychologism about Practical Reasons

There is a fundamental connection, in Williams's work, between what seems to be a "local" critique of some of the ideas of the Modern Western ethical tradition in the guise of the Morality System and a very general thesis about the nature of practical reasons in general. Unfortunately, Williams's thesis that all practical reasons ascriptions are "internal" has been subject to as much misunderstanding as any thesis he ever advanced. Even sophisticated moral philosophers persist in treating the view as a defense of a Humean instrumentalism about reasons combined with the implausible claim that all moral motivation is grounded on (unmotivated) desires. (This latter thesis is better described as Hobbesian rather than Humean.)

More sympathetic interpreters, however, such as Stephen Finlay (and the current author) trace Williams's thesis back to his interest in an Aristotelian conception of action explanation in general and Davidson's re-working of these Aristotelian ideas in particular [Davidson 1980; Thomas 2002, 2006; Finlay 2009]. As Finlay neatly encapsulates the view:

> [T]o believe that R is for you a reason for action *just is* to believe that R is a certain kind of explanation for action, an explanation of why you would act if you were to deliberate soundly. [Finlay 2009, pp. 1–2, emphasis added]

This expresses Williams's opposition to what he calls a "de-psychologised" conception of action explanation represented paradigmatically by the "external rea-

sons theorist." The external reasons theorist severs this connection between justification and even potential explanation.

Only this conception of reasons ascriptions as internal, Williams argued, captures *what it is* for a reason to be practical: reasons explanations must say something distinctive about the particular agent and explain why she acted as she did. How does this very abstract thesis about the explanation of rational agency bear on the very specific claims that Williams makes about the Morality System?

John Skorupski has argued that the connection emerges when one takes seriously Williams's claim that Kant was the "limiting case" of an internal reasons theorist: it is integral to our modern conception of morality that reasons be endorsable from the first-personal perspective of the agent. This is the idea that, as Skorupski puts it, "agents cannot be said to have reasons for acting which they are unable to recognise *as* reasons (even when they know the relevant facts)" [Skorupski 2007, p. 73]. If there is to be a threat to the Morality System, it cannot emerge from that thesis as both Williams and Kant share it.

The way in which the internal reasons thesis combined with this idea that reasons have to be endorsed from the perspective of the agent gives rise to a challenge to our moral outlook – at least as the Morality System represents it – emerges when it is combined with two other theses. The first is that of the *universality* of moral reasons. The second is a psychologically realistic account of actual agents and the variation in their psychological capacities. Skorupski gives the following example of a person, Tom, who lacks the capacity to feel gratitude:

> Imagine that Tom simply has no sense of gratitude.... So when Mary goes out of her way to help him, it's not possible that he should thank her for *that* reason, that is, simply and solely because he sees for himself that gratitude is appropriate. (He may of course recognise prudential reasons to observe the social conventions he's been told about, etc.) Does it follow that Tom does not *have* that reason for thanking Mary – that that particular fact is *not* a reason for him to thank her? I think our response to this kind of question is interestingly uncertain. [Skorupski 2007, p. 88]

Given Tom's incapacity, does he lack the relevant reason? If a reason is to play its explanatory role for a particular agent, then we must understand him or her as possessing the capacity to recognise it as a reason: our attitude to examples like Tom, Skorupski suggests, is indeed ambivalent.

But a commitment to the universality of moral reasons, plus the thesis that to blame a person requires that they be at fault for overlooking a reason available to them such that they could have done otherwise had they acknowledged it,

rules out this ambivalence. Hence the tension in our ordinary commitments – as we have historically inherited them.

Williams's conclusion is that we need to be more truthful about how our practices of praise and blame actually work. There is no such thing as a reasons ascription that can bear an "external" interpretation, but there are "optimistic internal reasons statements" that we direct at people in order to *make it true* that they have the very reasons on the basis of which we seek to blame them. Williams fully accepts the connection between his "cognitive internalism," as Skorupski calls it, and the grounding of blame on a reason potentially available to an agent:

> Blame rests, in part, on a fiction; the idea that ethical reasons, in particular the special kind of ethical reasons that are obligations, must, really, be available to the blamed agent.... *He ought to have done it*, as moral blame uses that phrase, implies *there was a reason for him to do it*, and this certainly intends more than the thought that we had a reason to want him to do it. [Williams 1995b, p. 16]

However, Williams immediately continues:

> But this may well be untrue: it was not in fact a reason for him, or not enough of a reason. Under this fiction, a continuous attempt is made to recruit people into a deliberative community that shares ethical reasons ... by means of this fiction people may well indeed be recruited into that community or kept within it. But the device can do this only because it is understood not as a device, but as connected with justification and with reasons that an agent might have had; and it can be understood in this way only because, much of the time, it is connected with those things. [Williams 1995b, p. 16]

It is not my brief here to assess, specifically, this aspect of Williams's critique of the Morality System.[6] Its bearing on my discussion is that Williams explains this "proleptic mechanism" for expanding the scope of blame (while not recovering its full universality) by invoking a reflexive motivation, on the part of most ethical agents, to "desire to be respected by people whom, in turn, one respects" [Williams, 1989, p. 7]. It is important to add that, for Williams, blame only works in the way it actually does if it is not privileged in the way the Morality System insists, but depends on other ethical dispositions. Furthermore, it is as subject as the internal reasons constraint – which it mirrors – to the constraint that, as a matter of contingent fact, people either have this "ethically important disposition" or they do not [Williams, 1989, p. 7].

---

6 For contrasting discussions of Williams's critique of the Morality System see Charles Taylor [1995] and Robert B. Louden [2007].

Given what Williams takes to be the most general truth about reasons there can, literally, be nothing that makes an external reasons ascription true. If blame is connected to such reasons in the way that the Morality System insists, then blame cannot be all that it seems to be. But *something* survives Williams's critical account in order to meet his test of stability under reflection: a conception of blame, in a more ramified set of connections with other ethical dispositions and emotions, and a truthful recognition of putatively external reasons as optimistic internal reasons statements. A realistic, unmoralised, account of human psychology leaves us with a conception of the ethical community as constituted by those who want to relate to others on a basis of mutual respect. We can truthfully live with each other on those terms and dispense with the fictions of the Morality System.

Similarly, Williams's later account of ground projects involves a generalization of an analogous reflexive structure to that involved in respect to those virtues that surround our practices of truth telling. But this more developed account starts even further back in Williams's account of "steadying the mind" and the constitutive involvement of a community in stabilizing the very disposition of belief such that the content entertained in an agent's mind can be, determinately, either a belief or a desire [Williams 2002, pp. 82–83]. This is a very ambitious argument that begins from Williams's famous claim that belief cannot be subject to the will, via an account of "steadying the mind" in the context of a community, to the way in which different conceptions of community support different conceptions of what it is to be truthful about oneself. However, Williams opts for *one* of these ideals of truthful self-expression as the correct one. His later account of ground projects features in his defense of this ideal [Williams 2002, chapter 9]. I will set out this argument in more detail in section four, below.

The starting point for any assessment of Williams's complex and mutually supporting set of arguments is this: in order for something to count as a reason for a particular agent, it must be potentially explanatory of that specific agent's actions as a result of sound deliberation from his or her initial subjective motivational set. Williams's moral psychology sees a certain kind of structuring identification – a ground project in Sartre's sense – as playing a distinctive functional role in any such set. As Pamela Hieronymi puts it: To appreciate that role we need to examine another aspect of Williams's ethical philosophy as a whole: his critique of act utilitarianism. Williams's constraints on the answer to the question, "Why be moral?" are substantive, not formal, and determine that the answer must fall within the class of what Scheffler calls "moderate moralities" [Scheffler 1993, p. 6].

> Now we can see how Williams' internalism underwrites his integrity objection .... [U]tilitarianism neglects the extent to which the agent's actions and decisions have to be seen as flowing from the projects and attitudes with which he is most clearly identified .... And yet utilitarianism tells the agents to treat his own projects and attitudes in an impartial way, to view them as if they belonged to someone else .... The agent can neither act nor deliberate from a truly impersonal standpoint, as "anyone"[7] [Hieronymi, n.d., pp. 4–5].

So I will now examine Williams's critique of act utilitarianism in more detail to show why his constraints on the answer to the question, "Why be moral?" are substantive, not formal. They determine that the answer to the question must fall within the class of what Scheffler calls "moderate moralities" [Scheffler, 1993, p. 4].

2 Integrity and the Rational Authority of Morality

It may seem to take some exegetical work to reconstruct the connection between ground projects and action from integrity, given that Williams' discussion of the latter idea occurs in a specific and apparently quite limited context: his critique of act utilitarianism. However, the appearance that Williams's point is merely local to that discussion is misleading:

> [T]he reason why utilitarianism cannot understand integrity is that it cannot coherently describe the relations between a man's projects and his actions. [Williams 1973, p. 100]

So the objection is the specific application of a more general thesis. I will suggest a phenomenological argument, implicit in Williams's discussion, which underwrites the claim that this argument has this wider significance.

In an important recent book Paul Hurley has given grounds for taking Williams's objection to act utilitarianism to have such significance for normative ethics [Hurley 2009]. In the guise of an objection to act utilitarianism, the point of Williams's argument is to highlight a fault line between the consequentialist's theory of value and theory of rightness, and between the latter and our intuitive idea of the rational authority of moral reasons. However, in the course of developing that argument Williams draws on a more general thesis: that to be

---

7 As Hieronymi notes here, this connects with Williams's thesis that practical reasoning is essentially first personal; this thesis was defended in Williams [1985] but later retracted – but Williams did not say why! For a different rationale for the thesis than Williams's own see Thomas [forthcoming].

identified with the ground project of Utilitarian morality is to be identified with a project with no distinctive content of its own. I will expand upon both points.

Hurley has argued that interpreters have misunderstood the nub of Williams's critique of utilitarianism. Standardly understood, it is a well-worn objection to act consequentialism that it tells us that, under certain circumstances, morality can be very demanding for an agent. One could, indeed, view the integrity objection through this lens: loss of integrity would feature as one more, admittedly heavy, cost to the agent. Unfortunately, however, the objection thus understood is not very plausible. It is, after all, not difficult to imagine circumstances in which the morally right action simply *is* very costly to an agent. Further, the act utilitarian can respond that one ought not to shoot the messenger: if act utilitarianism brings the message that *morality* can be very demanding, then that is a problem about the content of morality – and a practical problem for us. It is not, per se, a problem for act utilitarianism.

Hurley, however, argues that this is a misunderstanding: the real issue is not that act utilitarianism is too demanding, but that it makes no demands at all. The reason for this is that, qua species of the genus of consequentialist moral theories, this kind of normative theory is concerned solely with a constitutive account of the nature of rightness. The nature of rightness is explained in terms of the intrinsic values of outcomes. It says nothing at all about a separate relation, namely, about the relation between rightness and our intuitive notion of a reason. As Hurley aptly puts it, act utilitarianism "jettisons ordinary morality but leaves ordinary reason in place" [Hurley 2009, p. 183].

It simply takes the folk psychological conception of a moral reason for action for granted and fails to notice that the content that it gives to its substantive theory of rightness severs the connection between rightness and reasons. So far from issuing extreme demands, the theory (qua theory of rightness) issues none at all; it would only do so were it entitled to assume a connection between rightness and reasons. But it has no such entitlement: in fact, given the plausibility of Williams's project-based account of reasons for an agent, the theory gives us excellent advice as to how to *avoid* being alienated from the content of morality as the act utilitarian understands it:

> The real difficulty is not that rational agents are alienated *by* consequentialist morality from their plans, projects, and commitments, but that rational agents are alienated *from* such a consequentialist morality by the good reasons each has to honor her commitments and pursue her plans and projects. [Hurley 2009, p.21]

Consequentialism does not produce alienation from the content of morality, but the reverse. By the practical authority of moral reasons Hurley is appealing to no

more here than the minimal idea that agents can have sufficient or decisive reason to do what they ought to do [Hurley, 2009, p. 11–12].

If reasons are internal in Williams's sense, then that places him within the general family of "practical objectivity" theories that incorporate this minimal understanding of the authority of reasons (this is the "cognitive internalism" that Skorupski correctly attributes to him). So if internal reasons are grounded in an agent's ground projects, then act utilitarianism is powerless to alienate the agent from them. This focus on the distinct relations between the values of outcomes and rightness, and the relation between rightness and reasons, immediately explains the connection between these ideas and integrity: the act utilitarian tells an agent to perform the right action even if the agent has no reason at all to perform the recommended action. The act utilitarian requires acting rightly even if all the agent's internal reasons, based on her ground projects, give her most reason to do what is, by act utilitarian lights, the wrong action. That seems, in a very straightforward sense, a violation of an agent's integrity.

I am not concerned directly, here, with the assessment of this intriguing general argument against act utilitarianism, but with the deeper connection it makes between the internal reasons thesis and the nature of ground projects. As Hurley points out, in the standard understanding of the problem of moral demandingness both act utilitarians and their critics presuppose the rational authority of moral reasons. However, the very same act utilitarian account both *presupposes* that there are non-impersonal reasons and that agents have reasons to act on them. (That is what generates the initial paradox that agents have sufficient reasons to do wrong and insufficient reasons to do right.) But further reflection reveals how act utilitarianism, in it standard formulation, is self-stultifying – an even more serious problem.

That which produces this paradox is the claim that act utilitarianism must both exclusively represent the claims of the moral point of view *and* be rationally inescapable for an agent. Therefore, as an impersonal theory, the view cannot tolerate any non-impersonal reasons in its formulation. But Williams demonstrates that the theory rests essentially on non-impersonal reasons for its formulation such that, were it to acknowledge this truth, act utilitarianism would be merely one standpoint amongst other standpoints that the agent rationally could adopt. That latter concession is fatal to the theory; as Hurley's argument develops, it leads at least beyond consequentialism to hybrid theory and from there to a contractualism based on second-personal normativity. My interest here is in the first step of this argument: how act utilitarianism depends essentially in its formulation on non-impersonal reasons that it cannot acknowledge on pain of undermining its own claim to exclusive rational authority.

The crux of Hurley's reconstruction of the argument runs as follows: imagine an agent who takes upon herself the empirical equivalent of the normative task of Hare's World Agent: she represents the perspective of all other agents and their own interests, values and projects and ranks outcomes according to the maximal satisfaction of an act utilitarian standard.[8] [Hare 1981] There are no reasons in this picture as it is not a normative picture; the problem emerges when this conception is understood as potentially reason-giving with regard to ordinary agents. Under what conditions could we understand this empirical ranking as converted into a normative one? How could we understand the normative perspective of the World Agent?

Williams argues we can understand it only on one condition: that the agent who thus takes act utilitarianism as her ground project must be an exceptional agent – an agent whose position cannot be generalized to that of all others. If, however, she must be an exceptional agent then her standpoint loses its claim to be the exclusive standpoint in which the rational point of view and the moral point of view are unified. The standpoint of the World Agent is, then, a rationally *optional* standpoint and at that point the theory is in trouble. If it acknowledges that it rests on non-impersonal reasons, then it loses its claim to exclusive rational authority; if it fails to acknowledge that it is so grounded, then it denies an independently plausible phenomenological point about ground projects as part of its general severing of the relation between rightness and reasons.

What is the explanation of this necessarily exceptional status? Why is the standpoint of our putative utilitarian agent, for whom act utilitarianism is her ground project, one that cannot generalize to all other agents? The explanation Hurley recovers from Williams is that we *can* envisage an agent adopting act utilitarianism as her ground project. The problem blocking the generalization of this concession, however, is that this "higher order" ground project has no distinctive content of its own. The phrase "higher order" grants it neither any special content nor any special rational authority. Instead, it is one ground project amongst all other such projects.

Just as first order projects are transparent to the values that ground them, so the act utilitarian's higher order project of seeking compossibly to realize all the lower order projects of all agents (including her own) is transparent to those very same grounding values. Certainly the act utilitarian can stipulate that her exceptional ground project encompasses all others within its scope, but that is not to claim that being "higher order" captures any special rational authority (except

8 Hurley does not introduce the analogy with Hare's World Agent, but I take it his discussion suggests a view of this kind [Hurley, 2009, pp.70 – 71].

stipulatively). Comprehensive scope does not generate an automatic entitlement to such authority. Other agents are in a different position: they are free to adopt the act utilitarian's ground project and for them it could generate reasons, but that is because it is merely one optional project amongst all other projects and is, therefore, an explicit disavowal of such special authority.

Hence the dilemma: either the act utilitarian, in the light of the constitutive connection between ground projects and reasons, breaks the connection between rightness and reasons, or she does not. If she breaks it, she violates an agent's integrity by giving her no reason to do right, or good reason to do wrong. If she does not break it, she must abandon the claim that act utilitarianism shows that the moral and the rational points of view are identical and inescapable for all agents.

Thus the reason the specific discussion of the failures of act utilitarianism generalizes beyond this context is that Williams's critique of the act utilitarian draws on a general truth in moral psychology: that to be committed to a ground project is to be open to a range of values and reasons that structure an agent's deliberative field. It is the phenomenological claim that the life of virtue has no distinctive content of its own; it is a quasi-perceptual sensitivity to the values that are relevant to that ground project [Thomas 2005, 2006]. It is this phenomenological point that grounds the crucial claim that the "higher order" ground project of the envisaged utilitarian agent is essentially parasitic on the content of the ground projects of non-utilitarian agents. We can envisage an act utilitarian agent from whom the normative perspective of the World Agent *is* her ground project, so that for her the connection between rightness and reasons has been guaranteed; but if this ground project has no content of its own, then if we imagine it generalized to all agents then we are playing the "shell game." At some point we have to discharge the obligation to explain the content of the utilitarian ground project. When we do, its claim exclusively to represent a rationally mandatory standpoint that unites the rational and the moral is undermined. We cannot all be act utilitarians all of the time; this is true even if some of us can be act utilitarians some of the time.

The bearing of this point on the central question with which this paper is concerned, "Why be moral?" is that any answer to that question can only be raised, and answered, in the light of the substantive content of an agent's ground projects. Scheffler concludes, in his own "reconciliationist" project of formulating a hybrid theory, that the connection between rightness and reasons can be preserved only in what he called a "moderate morality" [Scheffler 1993; 1994]. That is the whole point of phrasing the question of whether a life good for the agent and a moral life are potentially convergent: there is no answer to that question that is independent of the substantive content of an agent's projects.

Given that reasons are internal, and that ground projects underpin our ordinary notion of acting for a reason in the light of a social identification, plus the constraint of psychological realism, Williams is eliminating competing positions one by one. The answer to the question, "Why be moral?" cannot involve a priori constraints on practical agency as such: in that sense it has no *formal* answer. Furthermore, in so far as the question has only a substantive answer, those answers are also constrained to fall within the class of "moderate moralities." It also seems we are in a position to say even more: that within that class we are constrained to consider those moderate moralities consonant with a background conception of ethical community that involves relations of mutual respect.[9] If even conceptualizing our ethical lives in that way represents a certain kind of historical achievement, can Williams's later discussion of political psychology determine the class of permissible answers to our initial question even further by turning from mutual respect to mutual acknowledgement?

3 From Moral Psychology to Political Psychology

I have outlined Williams's general strategy, but I have not yet allayed the concern that it fails to address the crucial issue. It seems that we need, case by case, to examine the substantive content of an agent's projects to see whether or not they ground a moral life satisfying to the agent. But is this eudaimonistic perspective sufficient to determine the content of *moral* projects? In order to address this point I think it is helpful to discuss Williams's final statement of his moral psychology.

The subject matter has changed, by the time of *Truth and Truthfulness*, to whether or not a genealogy of the concept of knowledge can lead us to identify virtues, norms, and values that are constitutively bound up with the "belief-assertion-communication system" [Williams 2002, p. 84; Thomas 2008]. Williams's answer is "yes," in an ambitious argument that tries to identify the schematic content of this system, based on interpretational arguments taken over from Donald Davidson, while also arguing that this schema has to be filled in via the contingent history of actual epistemic practices.

In order to highlight that his two key terms are terms of art, Williams capitalizes the words "Accuracy" and "Sincerity" as part of his general account of truth

9 My own view, which I will not develop at any length here, is that what I have called Williams's "phenomenological" argument pushes his view towards a form of moral realism – although this is not a conclusion that Williams would have welcomed. That Williams's resistance to realism partly depends on his epistemological foundationalism is argued in Thomas [2006/2010].

telling as involving these twin schematic dispositions on the part of people within the "system" of knowledge *to take care* and *not to lie* or otherwise mislead. It is not my aim here to assess the plausibility of Williams's characterization of these schematic epistemic virtues, but to focus on the way in which they shape Williams's final discussion of the idea of a ground project.[10]

Williams thought that, historically, different conceptions of Sincerity and Authenticity had played an important – sometimes fateful – role in shaping political expression of the social conditions to which these conceptions could be reconciled. Examining this argument in some detail reinforces the point that, for Williams, ground projects are not egoistic in content. This is brought out very clearly when the context of the discussion is the transition to the political expression of the ground project of being truthful about oneself.

In a suggestive discussion of the two contrasting accounts of Sincerity found in Rousseau and Diderot, Williams takes them to be two "elaborations" of his genealogy of Sincerity into two competing accounts of personal *authenticity* that explain "what it takes to be a truthful person" [Williams 2002, p. 173]. It is clear from his discussion that they mark two polar opposites – two ways of being wrong – but there is a correct account of the politics of self-knowledge recoverable from the insights of each view by a process of triangulation.

Williams claims that Rousseau's account of community is deeply monological in its account of truthfulness to self and *hence* truthfulness to others (in that order of explanation). Given that he believes he has direct knowledge of himself from immediate acquaintance, "le sentiment de l'existence," the mystery for Rousseau is what blocks his truthful presentation to others such that he is so persistently and frustratingly misunderstood[11] [Rousseau 1959, p. 1047]. The answer, for him, has to lie in extrinsic social conditions – when, that is, it is not simply a case of individual malice towards him.

So Rousseau further concludes that we need, collectively, to remove the social conditions of material inequality and hence dominance: we thereby remove the obstacles to mutual sincerity. Given that, for Rousseau, living with each other in conditions of spontaneous openness is the highest virtue, we need to live in a kind of political community where the blockages to this spontaneous openness have been removed.[12] This puts us in a position to institute a Rousseauian con-

10 I assess these arguments elsewhere in Thomas [2008].

11 "Why is it so difficult to bring about a concord between what one is for oneself and what one is for others?"[Starobinski, 1971/1988].

12 Williams quotes Starobinski on Rousseau's ideal for authentic inter-personal relations: "paradise was the reciprocal transparency of consciousnesses, a total and confident communication between them" [Starobinski, 1971/1988, p. 32 quoted by Williams, 2002, p. 174].

tract that involves no bargain, or trade, hence no cost to the agent: after signing the contract you are as free as you were before. We can see this as grounded in the deeper fact that "each, by giving himself to all, gives himself to no one".[13] [Rousseau, 1997, I.6.8] Primarily, we need to remove the obstacles to this openness; to remove the social causes of inflamed amour-propre that Williams revealingly describes as "competitive self-assertion."[14] Hence a republic that is restricted in size, where people are above all citizens at the expense of their private ends, renounce luxury, embrace frugality, and avoid all status inequalities that arise from the twin facts of material inter-dependence and inequality. From a liberal perspective, the problem is that the cost of this spontaneous openness if very high: this society is extensively coercive and the attractions of this modern day Sparta are elusive from our own perspective.

I turn now to Diderot who, by contrast, represents for Williams the claim that sociality enters into the stability of the dispositions that render the mind steady. This has the consequence that society is not merely the mutual collision of monologic truth-tellers as Rousseau envisages it. On Williams's reading of *Rameau's Nephew*, that work is concerned to present a character who is committed to Rousseauian transparency in his self-presentation, with the flaw that the self thus presented is not stable over time and thereby frustrates the goal of explaining what it is to be a truthful person. Diderot's performance model undoes itself: there is "uninhibited spontaneity," but with no grounding in a stable self. The protagonist improvises the self that his audience, in some specific context, requires him to be. (Perhaps the best analogy for this character is that of the conversational "shape shifter".)[15] So one virtue, spontaneous openness, does not secure another: mutual trust. It fails to do so as there is no stable character in whom one can trust. But the problem, as Williams sees it, actually runs deeper:

> On Diderot's view, as I understand him, it is a universal truth, not just a special feature of modernity, that human beings have an inconstant mental constitution that needs to be steadied by society and interaction with other people. [Williams 2002, p. 191]

13 As Philip Pettit notes, crucial here is the thought that no law legislated by the General Will can be unjust because "no man can be unjust towards himself" [Rousseau, 1997, II.6.7 quoted in Pettit, 2012, p. 15].

14 Taylor emphasizes Rousseau's claim that in this competition all lose: given a condition of dependence, "maître et esclave se dépravent mutuellment" [*Emile*, Paris, Garnier, 1964, Bk 2, p. 70, quoted in Taylor 1995].

15 There is a latter presentation of the same theme, of comparable brilliance to that of Diderot, in Melville's *The Confidence Man* [Melville, 1857/1991].

Williams's argument is that we cannot rest with the idea that Rameau's nephew is spontaneously authentic, but just for a moment; for a speaker of that kind whose character seems improvised and episodic, we ought to question "what kind of thing *is* in his mind" [Williams 2002, p. 191, emphasis added]. Episodic beliefs that change, not in response to evidence, but because of re-configurations of the agent's psychology, may come to strike us as not beliefs at all. This may not be the case with the deliverances of basic capacities of mind such as memory or perception. But for more complex interpretative or evaluative beliefs, these must be embedded in a social practice of situated avowal in order to be candidates for *belief* at all.

Paradoxically, Williams argues Diderot's insight allows for a more robust individualism than that envisaged by Rousseau. That is because of its recognition of "[the] social dimension [of] the construction of beliefs, attitudes and desires" which are "the materials of idiosyncrasy, and the lesson is that *we need each other in order to be anybody*" [Williams 2002, p. 200, emphasis added].[16] So we can now triangulate between the flawed views of Rousseau and Diderot: sociality enters into individuality and indeed makes it possible in a community characterized by various levels of inter-personal trust. Our fundamental relation to each other is dialogic and not monologic. This claim is grounded on fundamental facts in the philosophy of mind about knowledge, self-knowledge, and the stability of the dispositions that underpin belief.

Williams believed that a critical engagement with Rousseau and Diderot takes us directly to an argument for a liberal politics. Such a politics is a working out of the relationship between freedom from domination and an individual life worth living.[17] Rousseau was primarily concerned with freedom from domina-

16 Perhaps we should add "anybody in particular" given that, in his careful and complementary analysis of Rousseau's political psychology, Charles Taylor notes that, in a community of virtue bound together by "perfectly balanced reciprocity", for Rousseau it follows that "caring about esteem … is compatible with freedom and social unity, because the society is one in which all the virtues will be esteemed equally and for the same [right] reasons." "Paradoxically," Taylor continues, "the bad other dependence goes along with separation and isolation; the good kind, which Rousseau doesn't call other-dependence at all, involves the unity of a common project, even a 'common self'" [Taylor 1995, pp. 48–49 quoting *Du contrat social*, p. 244]. Williams's comment is that the aspiration to "coincide with myself only if I coincide with others" is a "desperate assertion … mere fantasy, a triumph of the wish" [Williams 2002, p. 200].
17 In a key range of cases that Williams describes, an individual's commitment to truthfulness stands in an antagonistic relation to a political culture "which destroys and pollutes the truth" [Williams, 2002, p. 127]. This takes a particular form in the special development, within modern societies, of the institutions of modern science and the vocation of a scientist who values truthfulness for its own sake. It matters to Williams that the scientist faces a recalcitrant reality that is

tion, but failed to secure the right kind of political community; Diderot recognizes that sociality is the underpinning of individuality so he sets us on the right track for a politics for a modern, commercial, society. Such a society involves both the values of individuality and the basis for a stable social order grounded on generalized impersonal trust and mutual assurance. A secure individualism is grounded in the social conditions for social cooperation.[18]

Interestingly for current purposes, Williams then extends his basic argument in a novel direction: this truth about the politics of knowledge helps to explain a distinctively modern phenomenon, namely, the politics of recognition. Modern authenticity depends on an "individualized" identity that is, as Charles Taylor puts it, "particular to me, and that I *discover* in myself" [Taylor 1995b, p. 227 emphasis added]. A refusal to recognize an identity is a form of personal or cultural impoverishment. Identities, in this politically relevant sense, have to reflect a combination of finding and making. This is Williams's only explicit discussion of the political dimension of his idea of a "ground project."

Williams argues that, first-personally, we need to be identified with them by an act of commitment, but concedes that even that word seems too voluntaristic. The role of such social identifications is that they have to be, from the first-personal perspective, as much found as made. That has a direct consequence for how we present ourselves to others: as calling for "acknowledgement." We need to be mutually recognized as being in the position to make claims upon each other that can be reciprocated. Our relation to each other in this endeavor is active and not passive. Williams describes the core of his liberal political psychology as follows:

> In the social or political case, where the presence of other people is vital, sincerity helps to construct or to create truth. Drawn to bind myself to the others' shared values, to make my own beliefs and feelings steadier [to make them, at the limit, for the first time into beliefs], I become what with increasing steadiness I can sincerely profess: I become what I have sincerely declared to them, or perhaps I become my interpretation of their interpretation of

not the will of another agent; "that is a key to the sense of freedom that it can offer" [Williams, 2002, p. 145]. To be free, in such a case, is to be free from arbitrary domination by the will of another, but also to be free of deceit: the virtues of truth are set against the difficulty of discovery of truths that are both independent of the will of the inquirer, but also free from interference by other inquirers.

18 To invoke Rousseau once again, it is social cooperation of the kind that he identifies in his famous parable of whether or not a group of individuals should hunt for stag (requiring mutual trust and a degree of risk, but with a higher payoff) or hunt for stag (requiring no trust, no risk, but with a lower payoff wholly within the individual's control [Rousseau, 1997, II.9]. Williams's answer, throughout *Truth and Truthfulness*, is that the "assertion-belief-communication" system rests on mutual assurance – in Rousseau's terms, a Stag Hunt.

> what I have sincerely declared to them. The sense that I am contributing to this, that it is a project, fills out the idea that acknowledgement is more than a mere factual discovery, while at the same time the sense that there is a discovery involved is related to the need to resist fantasy in making sense of my beliefs and allegiances in this way. [Williams 2002, p. 204]

Being sincere with each other is not, as Rousseau envisaged, the transparent expression of a truth about oneself that is already established, but is partially constituted by the stable underwriting of the disposition of belief that depends on our social nature. (Nor is this the inauthenticity of seeking to be only *what one is in the eyes of others* – an inverse of Rousseauian mirroring.) Claims about truthfulness, the epistemic virtues surrounding propositional truth, lead to surprising claims in the philosophy of mind. Those theses, in turn, support a view about the nature of political community where an essential role is played by mutual recognition and the very sociality that makes individuality possible. The schematic political psychology of *Truth and Truthfulness* envisages a genealogical argument from the nature of the knowing subject to two political conditions: that we live a life where we are not arbitrarily dominated by other agents and where the conditions that remove domination leave room for individual lives worth living.

Williams clearly envisages, in this late discussion, ground projects as figuring in the explanation of the worthwhileness of a life. *Truth and Truthfulness* contains what you might call a chastened account of their importance: he thinks they are rife with the capacity for self-delusion and error and always risk "ethical and social disaster" [Williams 2002, p. 205]. But they remain a part of his view and I also take it to be important that this late discussion takes them to be projects that call for acknowledgement; they are not a form of egoistic self-assertion. At this point it might seem as if we have gone full circle: Williams's "cognitive internalism" about reasons, the connection between blame and moral reasons, and this very ambitious transcendental argument in *Truth and Truthfulness* seem to offer a secure grounding for a morality of reciprocal recognition. In fact, Williams is making common cause with a conception of normativity as second-personal normativity, a (sophisticated) approach to contractualism. That is, indeed, how Paul Hurley interprets him [Hurley 2009].

My own view is that this conclusion would be too quick. Williams thinks our own political history has contained damaging mistakes about the particular ideal of truthfulness and its social pre-conditions; as I have indicated, he thinks that those mistakes can be corrected and one account can be selected as the truth on this issue. This conclusion is, however, fully compatible with his earlier discussion of the implications of his internal reasons thesis. That view, too, in-

voked a mechanism by which we seek mutually to stabilise each other's responses, but as I noted above it appealed to only one aspect of the morality of mutual recognition – the desire for mutual respect. In that sense the later discussion in *Truth and Truthfulness* further articulates this schematic account for a wider range of interlocking dispositions: any ideal of truthfulness of self that we can reflectively endorse forces us to embrace the existence of the more complex mechanism of mutual reliance and the constitutive role of community in "stabilising the mind."

I take it, however, that the transition from Williams's discussion of the proleptic mechanism of blame and its necessary fictions, even when expanded into the latter account of the community stabilisation of epistemic virtue, continues to draw on a reflexive truthfulness about these fictions. Williams thought that we had glossed over the working of the proleptic mechanism because the Morality System forces upon us the fiction that those we blame always had external reasons to act as they did such as to deserve blame. Recognise the fiction, however, and the device no longer functions as it is supposed to (just as, in *Truth and Truthfulness*, the stabilisation of the virtues of Sincerity and Accuracy requires them to be valued for their own sake or they will not discharge their function).

The two accounts now diverge: Williams's genealogy of Sincerity and Accuracy is compatible with a reflective understanding of them as more than mere devices, but and as bound up with two kinds of value valued for their own sake and a surrounding set of dispositions and emotions. The reflective account of the proleptic mechanism involved in blame, however, is stable under reflection only if, acknowledging the constraint of psychological realism, we have to abandon our presupposition that moral reasons are universal. What both accounts share, early and late, is that they are examples of Williams's wider methodological commitment to a naturalistic psychology for morality that does not presuppose any ethical ideas. The clearest expression of this goal is in *Shame and Necessity*, where Williams asks us to rethink the goals of moral psychology:

> This is not merely the idea that there must be a psychology of ethics .... The idea is rather that the functions of the mind, above all with regard to action, are defined in categories that get their significance from ethics. That is an idea that is certainly lacking in Homer and the tragedians. It was left to later Greek thought to invent it, and it has scarcely gone away from us since. [Williams 2008, p. 42]

It is this constraint that continues to make Williams, within the broad school of "mutual recognition" theorists, an outlier more indebted to Nietzsche than to any other of his historical forebears. It also seems to me to make him an uncomfortable bedfellow for proponents of a Fichte-inspired ideal of second-personal normativity, but a further consideration of that point awaits another occasion. At the

least it reinforces the point that the very idea of a contractualist account of morality is an internally complex one. It contains strands more hospitable to the very idea of internal reasons and those that are less accommodating [Freeman 1990; Wallace 1990; Thomas 2006/2010, chapter four].

Overall, then, how conclusive is Williams's discussion? We have conclusive reason to endorse the claim that any acceptable moral view will fall within the class of moderate moralities and we can rule out any a priori transcendental argument that tries to prove that morality is an inescapable ground project for any agent. In addition we can add the claim that any such ground project will seek acknowledgement from others just as, from within the moral properly understood freed from the distortion of the Morality System, moral reasons are exchanged between those who desire the respect of each other on a basis of mutuality.

The continuing causes for concern for Williams's critics are these: first, who gets to set the terms of acknowledgement?[19] Can the agent be selective in those from who she seeks acknowledgement? It is no help if the principled Nazi seeks, on her own terms, only acknowledgement from other Nazis. Point taken, but this simply raises the issue about content dependence again: the answer to our question requires a case-by-case consideration of substantive answers.

Secondly, with the necessary, a priori, and hence universal foundation for morality undermined, plus a constraint of psychological realism, what now guarantees the universality of moral reasons within such networks of mutual respect as those involved in Williams's proleptic mechanism? I submit that, if the perfect coincidence of a community of mutual respect and an evaluative community of mutual acknowledgement is no longer guaranteed by reason, then their relationship is, indeed, open to further determination.[20] Williams has identified constraints on the answer to the question, "Why be moral?" that are necessary, but not sufficient, conditions for an ethically adequate answer. But if Williams leaves us with a question, but not an answer, then it could hardly be otherwise. There are determinate truths to be had about the social and political conditions of any such answer and, as I have demonstrated, Williams establishes what they are in his discussion of the modern politics of identity. But the actual answer to

19 I am grateful to Brad Cokelet, Katerina Deligorgi and David Owen for separately raising this concern in different ways.

20 This is, in fact, the major choice point in contemporary contractualism between Scanlon's view, which presupposes an independently valuable idea of an ethical community bound by relations of mutual acknowledgement, and Darwall's reasons based view. I discuss this in 'Second Personal Reasons and Agent-Relativity: A Response to Mark LeBar' [Thomas, 2009b].

the question of "why be moral?" cannot be legislated for by philosophical reflection alone.

Conclusion

In this paper I have addressed Williams's answer to the question, "Why be moral?" by placing some of the fundamental concepts of his ethical outlook – the internality of reasons, the idea of a ground project, and the idea of action from integrity – in the context of his later political psychology. One, negative, aim is to address the worry that Williams's account of the place of ethical constraints in the life of an individual licenses egoism or, indeed, the expression "with integrity" of unethical motivations (such as those of the principled Nazi). There are prior, presuppositional, constraints in Williams's account, but it is not the set of a priori moral principles that the constitutivist argues constrain all expressions of practical agency. These constraints determine the answer to the question at different degrees of generality. Finally, any answer to the question for us – and, as usual in Williams's philosophy, identifying who "we" are is the crucial issue – is constrained by mutual acknowledgement and mutual respect, in the context of a modern ethical community where "modern" records the fact that a community of that kind is an historical achievement. I have argued that it is within the context of a community of this kind that candidate answers to the question, all within the range of moderate moralities, will be assessed not by philosophy but rather, as Alasdair MacIntyre once remarked, by reflective living [Thomas 2006/2010, p. 253].[21]

Bibliography

Davidson, Donald [1980] "Actions, Reasons and Causes," in *Essays on Actions and Events*, Oxford: Oxford University Press.

Diderot, Denis [2004] *Le neveu de Rameau* in *Contes et romans,* edited by Michel Delon, Paris: Collection Bibliothèque de la Pléiade.

Finlay, Stephen [2009] "The Obscurity of Internal Reasons," *Philosophers Imprint,* vol. 9, no. 7, pp. 1–22.

21 Research for this paper was funded by the Templeton Foundation research project "The Philosophy and Theology of Intellectual Humility." I am grateful to the Foundation for its support. For comments on earlier drafts of this paper I am grateful to Kathryn Brown, Brad Cokelet, Katerina Deligiorgi, Beatrix Himmelmann, Antti Kauppinen, Burt Louden, Kian Mintz-Woo, Adam Morton, and David Owen.

Freeman, Samuel [1990] "Reason and Agreement in Contract Views," *Philosophy and Public Affairs,* vol. 18, no. 2, pp. 122–157.

Hare, R. M. [1981] *Moral Thinking: Its Levels, Methods and Point,* Oxford: Oxford University Press.

Hieronymi, Pamela [Not dated] "Internal Reasons and the Integrity of Blame," unpublished manuscript.

Hurley, Paul [2009] *Beyond Consequentialism,* Oxford: Oxford University Press.

Korsgaard, Christine [2009] *Self-Constitution: Agency, Identity and Integrity*, Oxford: Oxford University Press.

Louden, Robert B. [2007] "The Critique of the Morality System" in Alan Thomas (ed.) [2007], pp. 104–134.

Melville, Herman [1857/1991] *The Confidence Man: His Masquerade*, Penguin Classics.

Nagel, Thomas [1970] *The Possibility of Altruism*, Oxford: Oxford University Press.

Pettit, Philip [2012] *On the People's Terms: A Republican Theory and Model of Democracy,* Cambridge University Press.

Ripstein, Arthur [2010] "Critical Notice: G. A. Cohen, *Rescuing Justice and Equality*," *Canadian Journal of Philosophy*, vol. 40, no. 4, pp. 669–700

Rousseau, Jean-Jacques [1959] *Les reveries du promeneur solitaire* in *Ouevres Completes,* Paris: Collection Bibliothèque de la Pléiade.

Rousseau, Jean-Jacques [1997] *Rousseau: 'The Social Contract' and Other Later Political Writings*, trans. Victor Gourevitch, Cambridge University Press.

Scheffler, Samuel [1993] *Human Morality,* Oxford: Oxford University Press.

Scheffler, Samuel [1994] *The Rejection of Consequentialism*, revised edition, Oxford: Oxford University Press

Skorupski, John [2007] "Internal Reasons and the Scope of Blame," in Alan Thomas (ed.) [2007], pp. 73–103.

Starobinksi, Jean [1988] *Jean-Jacques Rousseau: Transparency and Obstruction,* trans. Arthur Goldhammer, Chicago: University of Chicago Press.

Taylor, Charles [1995a] "A Most Peculiar Institution" in Altham and Harrison (eds.) [1995] *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge: Cambridge University Press, pp. 132–155.

Taylor, Charles [1995b] "The Politics of Recognition," in *Philosophical Arguments,* Cambridge, MA: Harvard University Press.

Thomas, Alan [2002] "Internal Reasons and Contractualist Impartiality," *Utilitas*, vol. 14, no. 2, pp. 135–154.

Thomas, Alan [2005] "Reasonable Partiality and the Agent's Point of View," *Ethical Theory and Moral Practice,* vol. 8, nos. 1–2, pp. 25–43.

Thomas, Alan [2006/2010] *Value and Context: The Nature of Moral and Political Knowledge*, Oxford: Oxford University Press.

Thomas, Alan (ed.) [2007] *Bernard Williams*, Cambridge: Cambridge University Press.

Thomas, Alan [2008] "The Genealogy of Epistemic Virtue Concepts," *Philosophical Papers*, vol. 37, no. 3, pp. 345–369.

Thomas, Alan [2009a] "Consequentialism, Integrity and Demandingness," in Timothy Chappell (ed.) *New Essays on the Problem of Moral Demandingness*, London: Macmillan.

Thomas, Alan [2009b] "Second Personal Reasons and Agent-Relativity: A Response to Mark LeBar" APA Pacific Division paper, April 11th, 2009.

Thomas, Alan [2015] "Rawls and Political Realism: Realistic Utopianism or Judgement in Bad Faith?," *European Journal of Political Theory* (forthcoming).

Thomas, Alan [forthcoming] "Practical Reasoning, the First Person, and Impartialism about Reasons" in Cottingham, J. and Feltham, B. (eds) *Partiality and Impartiality*, vol. 2, Oxford University Press.

Wallace, R. Jay [1990] "How to Argue about Practical Reason," *Mind*, vol. 99, no. 395, pp. 355–385.

Wallace, R. Jay [2001] "Normativity, Commitment and Instrumental Reason," *Philosophers Imprint,* December. Volume/ issue details?

Williams, Bernard [1973] "A Critique of Utilitarianism," in *Utilitarianism: For and Against*, with J. J. C. Smart, Cambridge: Cambridge University Press.

Williams, Bernard [1981] "Internal Reasons," in *Moral Luck*, Cambridge University Press.

Williams, Bernard [1989] "Internal Reasons and the Obscurity of Blame," *Logos*, 10, pp. 1–11.

Williams, Bernard [1995a] "Reply" in J. E. J. Altham and Ross Harrison (eds.) *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge: Cambridge University Press.

Williams, Bernard [1995b] *Making Sense of Humanity*, Cambridge Cambridge: University Press.

Williams, Bernard [2002] *Truth and Truthfulness: An Essay in Genealogy,* Princeton: Princeton University Press.

Williams, Bernard [2008] *Shame and Necessity,* Oakland: University of California Press.

Contributors

Dieter Birnbacher is Professor Emeritus of Philosophy at the University of Düsseldorf.

Richard Eldridge is Charles and Harriett Cox McDowell Professor of Philosophy at Swarthmore College, Pennsylvania.

Hallvard J. Fossheim is Associate Professor of Philosophy at the University of Bergen.

Roe Fremstedal is Professor of Philosophy at UiT, The Arctic University of Norway.

Beatrix Himmelmann is Professor of Philosophy at UiT, The Arctic University of Norway.

Christoph Horn is Professor of Philosophy at the University of Bonn.

Ivar Russøy Labukt is Associate Professor of Philosophy at UiT, The Arctic University of Norway.

Iddo Landau is Professor of Philosophy at the University of Haifa.

Robert B. Louden is Distinguished Professor of Philosophy at the University of Southern Maine.

Erik Lundestad is Associate Professor of Philosophy at UiT, The Arctic University of Norway.

Peter Schaber is Professor of Philosophy at the University of Zürich.

David G. Sussman is Professor of Philosophy at the University of Illinois at Urbana-Champaign.

Alan Thomas is Professor of Ethics at Tilburg University.

Héctor Wittwer is Professor of Philosophy at the Otto von Guericke University Magdeburg.

Index