



# SELF-CONSTITUTION

*Agency, Identity, and Integrity*

CHRISTINE M. KORSGAARD

OXFORD

## Self-Constitution

---

*This page intentionally left blank*

# Self-Constitution

---

*Agency, Identity, and Integrity*

Christine M. Korsgaard

OXFORD  
UNIVERSITY PRESS

# OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.

It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press  
in the UK and in certain other countries

Published in the United States

by Oxford University Press Inc., New York

© Christine M. Korsgaard 2009

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2009

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
without the prior permission in writing of Oxford University Press,  
or as expressly permitted by law, or under terms agreed with the appropriate  
reprographics rights organization. Enquiries concerning reproduction  
outside the scope of the above should be sent to the Rights Department,  
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover  
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

Typeset by Laserwords Private Limited, Chennai, India

Printed in Great Britain

on acid-free paper by

Antony Rowe Ltd, Chippenham

ISBN 978-0-19-955279-5 (Hbk)

978-0-19-955280-1 (Pbk)

10 9 8 7 6 5 4 3 2 1

*In Memory*

Marion (Maren) Hangaard Kortbek Korsgaard

1919–1999

*Perfect Mother, Best Friend*

*This page intentionally left blank*

# Contents

<i>Abbreviations for Frequently Cited Works</i>	ix
<i>Preface</i>	xi
1. Agency and Identity	1
1.1 Necessitation	1
1.2 Acts and Actions	8
1.3 Aristotle and Kant	14
1.4 Agency and Practical Identity	18
2. The Metaphysics of Normativity	27
2.1 Constitutive Standards	27
2.2 The Constitution of Life	35
2.3 In Defense of Teleology	37
2.4 The Paradox of Self-Constitution	41
3. Formal and Substantive Principles of Reason	45
3.1 Formal versus Substantive	45
3.2 Testing versus Weighing	49
3.3 Maximizing and Prudence	52
4. Practical Reason and the Unity of the Will	59
4.1 The Empiricist Account of Normativity	59
4.2 The Rationalist Account of Normativity	64
4.3 Kant on the Hypothetical Imperative	68
4.4 Against Particularistic Willing	72
4.5 Deciding and Predicting	76
5. Autonomy and Efficacy	81
5.1 The Function of Action	81
5.2 The Possibility of Agency	84
5.3 Non-Rational Action	90
5.4 Action	93
5.5 Attribution	100
5.6 The Psychology of Action	104



6. Expulsion from the Garden: The Transition to Humanity	109
6.1 Instinct, Emotion, Intelligence, and Reason	109
6.2 The Parts of the Soul	117
6.3 Inside or Outside?	122
6.4 Pull Yourself Together	125
7. The Constitutional Model	133
7.1 Two Models of the Soul	133
7.2 The City and the Soul	135
7.3 Platonic Virtues	142
7.4 Justice: Substantive, Procedural, and Platonic	148
7.5 Kant and the Constitutional Model	153
8. Defective Action	159
8.1 The Problem of Bad Action	159
8.2 Being Governed by the Wrong Law	161
8.3 Four or Five Bad Constitutions	165
8.4 Conceptions of Evil	170
8.5 Degrees of Action	174
9. Integrity and Interaction	177
9.1 Deciding to be Bad	177
9.2 The Ordinary Cases	180
9.3 Dealing with the Disunified	184
9.4 Kant's Theory of Interaction	188
9.5 My Reasons	197
9.6 Deciding to Treat Someone as an End in Himself	200
9.7 Interacting with Yourself	202
10. How to be a Person	207
10.1 What's Left of Me?	207
10.2 Conclusion	212
<i>Bibliography</i>	215
<i>Index</i>	219

# Abbreviations for Frequently Cited Works

References to and citations of frequently cited works are given parenthetically in the text, using the abbreviations cited below. For the editions and translations quoted, please see the Bibliography.

## 1. Aristotle

References to Aristotle's works will be given by the standard Bekker page, column, and line numbers, using the following abbreviations.

NE	<i>Nicomachean Ethics</i>
M	<i>Metaphysics</i>
MA	<i>Movement of Animals</i>
OS	<i>On the Soul</i>
POL	<i>Politics</i>

## 2. Hume

T     *A Treatise of Human Nature* (cited by book, section, and page number)

## 3. Kant

References to Kant's works will be given by the page numbers of the relevant volume of *Kants gesammelte Schriften*, which appear in the margins of most translations. The *Critique of Pure Reason*, however, is cited in its own standard way, by the page numbers of both the first (A) and second (B) editions. The abbreviations used follow:

ANTH	<i>Anthropology from a Pragmatic Point of View</i>
C1	<i>Critique of Pure Reason</i>
C2	<i>Critique of Practical Reason</i>
CBHH	"Conjectures on the Beginning of Human History"
G	<i>Groundwork of the Metaphysics of Morals</i>
IUH	"Idea for a Universal History with a Cosmopolitan Purpose"
LE	<i>Lectures on Ethics</i>

MM	<i>The Metaphysics of Morals</i>
PP	<i>Perpetual Peace</i>
REL	<i>Religion within the Limits of Reason Alone</i>

#### 4. Nietzsche

GM	<i>The Genealogy of Morals</i>
----	--------------------------------

#### 5. Plato

References to Plato's works are inserted into the text, using the standard Stephanus numbers inserted into the margins of most editions and translations of Plato's works. Other works are indicated by title.

R	<i>Republic</i>
---	-----------------

#### 6. My own works

CA	<i>The Constitution of Agency</i>
CKE	<i>Creating the Kingdom of Ends</i>
SN	<i>The Sources of Normativity</i> (cited by section and page number)

# Preface

One who is just does not allow any part of himself to do the work of another part or allow the various classes within him to meddle with each other. He regulates well what is really his own and rules himself. He puts himself in order, is his own friend, and harmonizes the three parts of himself like three limiting notes in a musical scale—high, low, and middle. He binds together those parts and any others there may be in between, and from having been many things he becomes entirely one, moderate and harmonious. Only then does he act.

(Plato, *Republic* 443d–e)

Both human beings and the other animals act, but human actions can be morally right or wrong, while the actions of the other animals cannot. This must be because of something distinctive about the nature of human action, about the way in which we human beings make choices. In this book, I try to explain what that distinctive feature is, and how it is connected to some of the other things that make human life different from the lives of the other animals. The name I give to the distinctive feature is the traditional one—rationality. As I understand it, reason is a power we have in virtue of a certain type of self-consciousness—consciousness of the grounds of our own beliefs and actions. This form of self-consciousness gives us a capacity to control and direct our beliefs and actions that the other animals lack, and makes us active in a way that they are not. But it also gives us a problem that the other animals do not face—the problem of deciding what to count as a reason for belief or action. To put the point another way, this form of self-consciousness makes it necessary to *take control* of our beliefs and actions, but we must then work out how to do that: we must find normative principles, laws, to govern what we believe and do. The distinctive feature of human beings, reason, is therefore the capacity for normative self-government.

The capacity for normative self-government brings with it another distinctively human attribute, normative self-conception, perhaps more than anything else the thing that makes being human both an adventure and a curse. For an action is a movement attributable to an agent as its author, and that means that whenever you choose an action—whenever you take control of your own movements—you are constituting yourself as the author of that action, and so you are deciding who to be. Human beings therefore have a

distinct form of identity, a norm-governed or practical form of identity, for which we are ourselves responsible. As a rational being, as a rational agent, you are faced with the task of *making something* of yourself, and you must regard yourself as a success or a failure insofar as you succeed or fail at this task.

If, when we act, we are trying to constitute ourselves as the authors of our own movements, and at the same time, we are making ourselves into the particular people who we are, then we may say that the function of action is *self-constitution*. This conception of action opens up the possibility that the specific form of goodness or badness that applies to human actions—rightness or wrongness—is goodness or badness *of their kind*, goodness or badness *as actions*. A good action is one that constitutes its agent as the autonomous and efficacious cause of her own movements. These properties correspond, respectively, to Kant's two imperatives of practical reason. Conformity to the categorical imperative renders us autonomous, and conformity to the hypothetical imperative renders us efficacious. These imperatives are therefore constitutive principles of action, principles to which we necessarily are trying to conform insofar as we are acting at all.

That way of putting it will make it clear that the conception of morality and practical reason that I defend in this book is the Kantian one. But I also draw on the work of Aristotle, to explain the sense in which an intentional movement can be attributed to an agent as its author, and on Plato, to explain the kind of unity that a person must have in order to be regarded as the author of her movements. For it is essential to the concept of an action that it is attributable to the person as a whole, as a unit, not to some force that is working in her or on her. And it was Plato who taught us, in the *Republic*, that the kind of unity required for agency is the kind of unity that a city has in virtue of having a just constitution.

Following Plato's lead, in this book I argue that the kind of unity that is necessary for action cannot be achieved without a commitment to morality. The task of self-constitution, which is simply the task of living a human life, places us in a relationship with ourselves—it means that we interact with ourselves. We make laws for ourselves, and those laws determine whether we constitute ourselves well or badly. And I argue that the only way in which you can constitute yourself well is by governing yourself in accordance with universal principles which you can will as laws for every rational being. It follows that you can't maintain the integrity you need in order to be an agent with your own identity on any terms short of morality itself. That doesn't mean that we have a reason for being moral that is selfish, that morality gets us something *else*, the integrity needed for agency and identity. Rather, it means

that a commitment to the moral law is built right into the activity that, by virtue of being human, we are necessarily engaged in: the activity of making something of ourselves. The moral law is the law of self-constitution, and as such, it is a constitutive principle of human life itself.

In this book I argue that thinking is just talking to yourself, and talking is just thinking in the company of others (9.4.12). I have been working on this book for a long time, and have had the benefit of thinking in the company of many others. It was as a result of reading Derek Parfit's work that I first began to think of identity as a problem that human beings have to solve (see my "Personal Identity and the Unity of Agency: A Kantian Response to Parfit," CKE essay 13). I am grateful to Derek for that, and I apologize to him for kidnapping his Russian nobleman for my own purposes (see Chapter 9). Jay Schleusener once remarked to me that Socrates was "good at being a person," and anyone who reads this book will see what I have made of that thought. I first published a short version of the ideas found here as an essay, "Self-Constitution in the Ethics of Plato and Kant," in 1999 (*Journal of Ethics*, 3: 1–29; now CA essay 3), after delivering it in various places as a talk. I would like to thank audiences at the inaugural meeting of the Society for Ethics at the Eastern Division Meetings of the American Philosophical Association, the University of Amsterdam, the University of Constance, the Humboldt University of Berlin, the University of Pittsburgh, the University of Virginia, the University of Salzburg, the University of Toronto, York University of Toronto, and the University of Zurich for discussions of that paper.

I presented some of the other ideas found here as a talk, "Human Action and the Kantian Imperatives," to audiences at a joint session of the University of Adelaide and Flinders University, at the University of Auckland, as the Simone Weil Lecture at Australian Catholic University, at the Autonomous University of Mexico, the University of Canterbury, the University of Dunedin, the Research School of Social Sciences in Canberra, and Stanford University. I am grateful to all of these audiences for helpful and stimulating discussion.

I expanded the longer manuscript from which those essays were drawn into a set of lectures, which I delivered as the Locke Lectures at Oxford in 2002. I am deeply grateful to members of the Department of Philosophy at the University of Oxford for that opportunity and for their hospitality while I was there. I also delivered those lectures as the Hägerström Lectures at Uppsala in 2002, and individual lectures from the set, over the next few years, at the Scots Philosophical Society at the University of Aberdeen, as the Paton Lecture at St Andrews University, at the University of California at Berkeley, at Brandeis University, as keynote speaker at a Brown University Graduate

Student Conference, at a conference at the Center for Subjectivity Research at the University of Copenhagen, as the Royal Institute of Philosophy Lecture at Durham University, at the University of Glasgow, at the Central States Philosophical Association at the University of Missouri at Columbia, at Ohio University, at the University of Reykjavik, at the University of Stirling, and at Union College. I am grateful to the audiences on all of these occasions.

I also had very helpful discussions of the manuscript with the faculty of the Department of Philosophy at Lehigh University, and with the Colloquium on Ethics, Politics, and Society at LUISS University in Rome. I am especially grateful to Susan Wolf and the members of her seminar in 2004 for their extremely helpful and thought-provoking questions. I had the benefit of a great deal of commentary at a conference devoted to my work in Madrid in 2004. I would especially like to thank Mary Clayton Coleman, Ana Marta González, and David Dick for their comments on that occasion.

For individual comments on the manuscript at various stages of its evolution I would like to thank Melissa Barry, Alyssa Bernstein, Charlotte Brown, Sarah Buss, Douglas Edwards, Barbara Herman, Govert den Hartogh, Peter Hylton, Arthur Kuflik, Anthony Laden, Jonathan Lear, Anton Leist, Richard Moran, Julius Moravcsik, Herlinde Pauer-Studer, David Plunkett, Andrews Reath, Miles Rind, Arthur Ripstein, Amélie Rorty, Tamar Schapiro, and Theo van Willigenburg. I am sure that this list is radically incomplete—indeed I know that it is, for I have in my files an extremely helpful set of written comments, unsigned, the author of which I am unable to identify. I can only apologize to everyone who I have left out.

The book was completed under the auspices of a grant from the Mellon Foundation, to whom I am profoundly grateful. I would also like to thank Eylem Özaltun, Nataliya Palatnik, and Paul Schofield for proofreading the manuscript, saving it from many errors, and suggesting many helpful changes. And finally I would like to thank Douglas Edwards for help selecting the cover, and Peter Momtchiloff for his apparently boundless patience during the many years it took me to get ready to publish the book.

*Christine M. Korsgaard*  
*May 2008*

# Agency and Identity

## 1.1 Necessitation

### 1.1.1

Human beings are *condemned* to choice and action. Maybe you think you can avoid it, by resolutely standing still, refusing to act, refusing to move. But it's no use, for that will be something you have chosen to do, and then you will have acted after all. Choosing not to act makes not acting a kind of action, makes it something that you do.

This is not to say that you cannot fail to act. Of course you can. You can fall asleep at the wheel, you can faint dead away, you can be paralyzed with terror, you can be helpless with pain, or grief can turn you to stone. And then you will fail to act. But you can't *undertake* to be in those conditions—if you did, you'd be faking, and what's more, you'd be acting, in a wonderfully double sense of that word.<sup>1</sup> So as long as you're in charge, so long as nothing happens to derail you, you must act. You have no choice but to choose, and to act on your choice.

So action is necessary. What kind of necessity is this? Philosophers like to distinguish between *logical* and *causal* necessity. But the necessity of action isn't either of those. There's no logical contradiction in the idea of a person not acting, at least on any particular occasion. You could not fail to act, in all the ways I've just described, if there were. And although particular actions, or anyway particular movements, may have causes, the general necessity of action is not an event that is caused. I'm not talking about something that works *on* you, whether you know it or not, like a cause: I am talking about a necessity you are *faced* with.

Now sometimes we also talk about *rational* necessity, the necessity of following the principles of reason. If you believe the premises, then you *must* draw the conclusion. If you will the end, then you *must* will the means. That's rational necessity, and it's a necessity you are faced with, so that comes closer.

<sup>1</sup> Later I will argue that, for essentially the same reason, you cannot undertake to be an unjust person (9.1.5).



But the necessity of action isn't quite like that either, for in those cases we have an if-clause, and the necessity of action is, by contrast, as Kant would say, unconditional. The necessity of choosing and acting is not causal, logical, or rational necessity. It is our *plight*: the simple inexorable fact of the human condition.

### 1.1.2

But once inside that fact, once we face the necessity of acting, we are confronted with a different kind of necessity. We live under the pressure of a vast assortment of laws, duties, obligations, expectations, demands, and rules, all telling us what to do. Some of these demands are no doubt illicit or imaginary—just social pressure, as we say (as if we knew what that was). But there are many laws and demands that we feel we really are bound to obey. And yet in many cases we would be hard pressed to identify the source of what I call the *normativity* of a law or a demand—the grounds of its authority and the psychological mechanisms of its enforcement, the way that it binds you. In philosophy we raise questions about the normativity of highbrow laws like those of moral obligation or theoretical and practical reason. But it is worth remembering that in everyday life the same sort of questions can be raised about the normativity of the laws and demands of professional obligation, filial obedience, sexual fidelity, personal loyalty, and everyday etiquette.<sup>2</sup> And just as we may find ourselves rebelling against, say, the sacrifice of our happiness to the demands of justice, so also, in a smaller, more everyday way, we may find ourselves *bucking* against doing our chores or returning unwanted phone calls or politely thanking a despised host for a dull party.

The surprising thing is not that we resist such demands, but that our resistance so often fails. Sometimes to our own pleasant surprise, sometimes merely with bewilderment or bemusement, we find ourselves doing what we think we ought to do, in the teeth of our own reluctance, and even though

<sup>2</sup> Some philosophers think this is not a separate point: they think that insofar as we are bound by the laws and demands listed here it is because they are morally obligatory. That view in turn can take several forms: (1) they are derivable from moral considerations; (2) although not derivable from moral considerations, all of their normativity (if they really have any) is derived from morality; (3) they are independent sources of normativity, but morality backs and reinforces them. The first of those options founders over the obvious cultural relativity and situational nature of the obligations in question: universal reason does not tell us, say, when to bring gifts. On the second view, there is something like a moral obligation to be polite, whose details are filled in by cultural practice. My own view, spelled out in *The Sources of Normativity*, is a version of the third option: they are independent sources of normativity, yet may require moral backing if they are to maintain their normative force in the face of reflection. (SN 3.4.7–3.6.1, pp. 120–30, and section 5 of the Reply, pp. 251–8.) I believe we should prefer the third view, since people in whose lives morality has only a tenuous footing may nevertheless experience and take seriously these lower-order demands.

nothing obvious forces us to do it. We toil out to vote in unpleasant weather, telephone relatives to whom we would prefer not to speak, attend suffocatingly boring meetings at work, and do all sorts of irksome things at the behest of our families and friends. Part of the lawless charm of a character like W. C. Fields springs from the fact that most of us are almost incapable of ignoring the requests of children, and yet we chafe under the enthrallment. It is a fact worthy of philosophical attention that the wanton disregard of life's little rules makes the people who would never break them laugh. To be sure, there is no question that in what Joseph Butler called "a cool hour," most of us would unhesitatingly choose to be the kinds of people who generally do what they ought.<sup>3</sup> As Aristotle observes:

no one would maintain that he is happy who has not in him a particle of courage or temperance or justice or practical wisdom, who is afraid of every insect which flutters past him, and will commit any crime, however great, in order to gratify his lust for meat or drink, who will sacrifice his dearest friend for the sake of half a farthing, and is as feeble and false in mind as a child or a madman. (POL 7.1 1323a26–34)

But there is also no question that in those warmer hours when we actually choose the particular actions demanded of us, we often manifestly do not *want* to do them. And yet we do them, all the same: the normativity of obligation is, among other things, a psychological force. Let me give this phenomenon a name, borrowed from Immanuel Kant. Since normativity is a form of necessity, Kant calls its operation within us—its manifestation as a psychological force—*necessitation*.<sup>4</sup>

### 1.1.3

In recent years, it has become rather unfashionable to focus on the phenomenon of necessitation. It seems to evoke the lugubrious image of the good human being as a Miserable Sinner in a state of eternal reform, who must constantly repress his unruly desires in order to conform to the demands of duty. Necessitation is thus conceived as *repression*. In opposition to this, some recent virtue theorists have offered us the (to my mind) equally rebarbative picture of the virtuous human being as a sort of Good Dog, whose desires and inclinations have been so perfectly trained that he always does what he ought to do spontaneously and with tail-wagging cheerfulness and enthusiasm. The opposition between these two pictures is shallow, for they share the basic

<sup>3</sup> See Joseph Butler, *Five Sermons*, ed. Stephen Darwall. The quotation is from Sermon 4 (Sermon 11 of the original *Fifteen Sermons Preached at the Rolls Chapel* from which these five are selected), paragraph 20, p. 56.

<sup>4</sup> *Nötigung* (G 4:413).

intuition that the experience of necessitation is a sign that there is something wrong with the person who undergoes it.<sup>5</sup> The disagreement is only about how inevitable the evil is. It may be natural to think of necessitation as a sign that something is wrong, since necessitation can be painful, and it is natural to interpret pain as a sign that something is wrong. But necessitation is so characteristic, so utterly commonplace a feature of human experience, that we should not be in a hurry to jump to that conclusion. In the *Republic*, Socrates says that the phrases we use to describe necessitation, phrases like “self-control” or “self-mastery” or “self-command,” seem absurd on their surface, since the stronger self who imposes the necessity is the same person as the weaker self on whom it is imposed. But Socrates also suggests that these phrases are like “tracks or clues” that virtue has left in the language (R 430e). Necessitation, he thinks, reveals something important about human nature, about the constitution of the human soul. What it reveals—that the source of normativity lies in the human project of self-constitution—is my subject in this book.

#### 1.1.4

The trouble with those two images of virtue—the Reformed Miserable Sinner and the Good Dog—and with the philosophical theories behind them, is not merely that they denigrate the experience of necessitation. It is also (and relatedly) that they do not give an adequate explanation of how we are necessitated. This is a somewhat complicated point, but let me try to explain what I have in mind.

Although there is reason to doubt whether David Hume would have accepted the characterization of normativity as a kind of necessity and its operation as a kind of necessitation, his theory of the natural virtues will serve my purposes well here, so I’m going to ask you to set those doubts aside. Just think about the question how normativity operates *in* us on the Humean view. Hume believes that moral concepts are generated from the point of view of a spectator, moved by sympathy to form sentiments of impartial love or hatred, inspired by people’s dispositions and motives, which we accordingly deem “virtues” or “vices.” For example, if I see that you pay your employees less than they need to live on, and reason that your desire for profit is the motive, then sympathy with your employees may lead me to condemn your desire for profit, and to call it the vice of greed. The person

<sup>5</sup> Kant himself seems to fall into this error in the *Groundwork*, when he suggests that necessitation is only experienced by an imperfectly rational will (G 4:414). I argue that this should not be Kant’s considered view in “The Normativity of Instrumental Reason” (CA essay 1), pp. 51–2.

Hume calls “naturally virtuous” is simply someone whose dispositions and motives are those that, from this spectator point of view, strike us as lovable. She is, for example, beneficent, and the spectator, sympathizing with those whom she benefits, finds her beneficence lovable and so judges it to be a virtue.<sup>6</sup> Although this standard of virtue is a human standard derived from a human point of view, and not an external standard imposed by God or Objective Values, it is nevertheless external to the operative dispositions and motives of the naturally virtuous person herself. The naturally virtuous person as such isn’t necessitated or even motivated by the standard of virtue: she is good, but not because the standard of virtue is a force that operates within her.<sup>7</sup> Nor is the goodness of her dispositions and motives in any essential way the result of her own possession of that standard; she might have cultivated her dispositions on purpose to meet the standard, but then again, she might not—she might just be *naturally* good. If we ask why the good person does *virtuous* actions *as such*, there is no real answer. Hume can only say: “well, that’s just what it *means* to be naturally good: it is to be the sort of person who has the sorts of dispositions and desires that spectators call ‘virtuous’.” Like Francis Hutcheson before him, Hume allows the standard of virtue to operate as a psychological force only in a second-best case.<sup>8</sup> He says:

When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least, to disguise to himself, as much as possible, his want of it. (T 3.2.1, 479)

So in the naturally virtuous person, the normative standard does not operate as a psychological force; and in this second-best case, it does so only through the medium of self-hatred. Normative standards are not, in and of themselves, psychological forces at all.<sup>9</sup>

Sentimentalist theories of this kind originally developed (and still usually situate themselves) in opposition to dogmatic rationalist theories, according

<sup>6</sup> Here I am summarizing Hume’s views as they appear in *A Treatise of Human Nature*, Book 3, part 3. The two examples I’ve given may seem different, since in the second the character trait in question (beneficence) seems to be identifiable prior to moral judgment, while in the first (greed), it does not. Strictly speaking I think Hume’s view should be that the character trait in question is not identifiable prior to moral judgment, although that makes the examples a bit cumbersome to spell out. On this see my “The General Point of View: Love and Moral Approval in Hume’s Ethics” (CA essay 9), pp. 295–6.

<sup>7</sup> It is somewhat different in the case of what Hume calls the “artificial” virtues, such as justice. The just person is motivated by thoughts about what is required by justice. But, as I mention below, it is arguable that those thoughts do not move him directly, but only because he has a desire to avoid his own self-hatred or self-disapproval.

<sup>8</sup> For Hutcheson’s view on this point, see his *Illustrations on the Moral Sense*, section 1.

<sup>9</sup> See Charlotte Brown, “Is Hume an Internalist?”

to which normativity—in particular, rightness—is an objective property grasped by reason. On such views, the rightness of an action, or for that matter the logical force of an argument, is an objective fact about the external world, which the rational mind as such grasps, and to which it then conforms its beliefs and actions. Dogmatic rationalists do suppose that normative standards operate within us as psychological forces, for reason may have to exert its force against unruly desires. But on reflection we can see that dogmatic rationalists transfer to *reason itself* the same bland—and seemingly blind—conformity to external standards that sentimentalists attribute to the dispositions of the naturally virtuous person. Dogmatic rationalists believe that norms exist outside of human reason—they arise from Objective Values or Moral Facts or some sort of rational structure that exists “out there” in the universe. But if reflection on that fact prompts us to ask *why* human reason finds it necessary to conform to these standards, there is no real answer. The dogmatic rationalist can only reply: “well, that’s just what it *means* to be rational, to have a mind (or a will) that conforms to the standards that we call ‘rational’.” In fact in theories of this kind “human reason” is really nothing more than the name of that faculty within us, whatever it might be, that conforms to rational standards. It is not identified in any other way.<sup>10</sup> Reason as envisaged by these theories is like a normative module that has been inserted into you, for the purpose of making the laws of reason, which are essentially outside of you, also be a force within you. Human reason is Objective Reason’s little representative within. And if we ask what gives rise to the psychological necessity of conforming to the laws of reason, the answer is in effect just to *point* at the module: human beings have a representative of reason within them and *that* makes it necessary. But, really, why should *we* conform to the demands of this little representative within us, or for that matter, why does *it* conform so readily to the demands of Objective Reason outside? Like other homuncular theories, dogmatic rationalism does not give an explanation of reason’s capacity to bind us, but merely points with premature satisfaction at the place where the explanation must go.

So there’s a parallel here: if we ask Hume why the good person conforms to the standard of virtue, there is no answer: we can only say, “that’s just what it means to be a good person.” And if we ask the dogmatic rationalist why human reason conforms to the standards of reason, there is again no answer: we can only say, “that’s just what it means to be reason.” It is certainly

---

<sup>10</sup> For other discussions of this point see my Introduction to *The Constitution of Agency*, pp. 2–3; “The Normativity of Instrumental Reason” (CA essay 1), pp. 55–6; and “Acting for a Reason” (CA essay 7), pp. 212–15.

true that from a third-personal point of view, when we call people vicious or irrational, we mean that they fail to conform to certain standards. But that failure is the outward manifestation of an inner condition, and these theories do not tell us what that inner condition is. They don't tell us how we are necessitated.

### 1.1.5

Let me try to phrase my complaint in a slightly different way. According to the dogmatic rationalist theories I've just described, when we do experience necessitation, it is the necessitation involved in the struggle *to act rationally*. The *goal* of acting rationally itself is taken for granted: but there are forces within us, unruly desires, which sometimes interfere with our capacity to meet that goal. It is only then that we must exercise self-command, and that we experience necessitation. Otherwise, being rational would be effortless. In the same way, in Hume's account of natural virtue, necessitation (or something like it) only occurs in the second-best agent who is explicitly trying *to be good*. In the best case, in the naturally virtuous person, goodness is effortless. So according to these theories, insofar as the standards of goodness or rationality operate within us as psychological forces, they operate as unachieved or endangered *goals*.

I believe that these theories both underestimate and misplace the role of necessitation in our psychic lives. There is work and effort—a kind of struggle—involved in the moral life, and those who struggle successfully are the ones whom we call “rational” or “good.” But it is not the struggle *to be rational* or *to be good*. It is, instead, the ongoing struggle for integrity, the struggle for psychic unity, the struggle to be, in the face of psychic complexity, a single unified agent. Normative standards—as I am about to argue—are the principles by which we achieve the psychic unity that makes agency possible. The work of achieving psychic unity, the work that we experience as necessitation, is what I am going to call *self-constitution*.

### 1.1.6

More specifically, in this book I will be dealing with three topics that I take to be intimately related. The topics are the nature of action, the constitution of personal or practical identity, and the normativity of the principles of practical reason.<sup>11</sup> For the sake of orientation I am going to begin by laying out the

<sup>11</sup> For an argument in favor of seeing the issue of personal identity in practical terms, see my “Personal Identity and the Unity of Agency: A Kantian Response to Parfit” (CKE essay 13).

basic elements of the conception that I believe relates these topics. Necessarily, what I say at this early stage will seem mysterious and cryptic, or at the very least dogmatic. So I ask you to keep in mind that this is a summary of a view to be defended in detail in the rest of the book and that nothing that follows is meant to be uncontroversial or obvious.

## 1.2 Acts and Actions<sup>12</sup>

### 1.2.1

Let me begin with the nature of action. If we want to learn what it is that makes actions right or wrong, we must start by asking what actions are, what their function is (2.1.1). John Stuart Mill thought he knew the answer to both of these questions. In the opening remarks of *Utilitarianism*, he says:

All action is for the sake of some end, and rules of action, it seems natural to suppose, must take their whole character and color from the end to which they are subservient.<sup>13</sup>

According to Mill, action is essentially production, and accordingly its function is to bring something about. Whether an action is good depends on whether *what* it brings about is good, or as good as it can be.

The influence of this conception of action on contemporary Anglo-American moral philosophy has been profound. Nowadays even moral philosophers who are not utilitarians appear to be comfortable only if they can explain moral value in terms of the production of various goods and harms. Deontological considerations are sometimes characterized as “side-constraints,” as if they were essentially restrictions on ways to realize ends.<sup>14</sup> As such, they have been found mysterious by many philosophers.<sup>15</sup> If the whole point of action is to produce the good, how then can it be good to restrict that production? A standard move in utilitarian arguments, a move that Jeremy Bentham made right from the start, is to insist that productive success—effectiveness for good—is an obvious, unquestionable standard for actions. The burden of proof, he argues, is on his opponents to show that there is any other standard that actions have to meet.<sup>16</sup> And while many moral

<sup>12</sup> Some of the material in this section has appeared in “From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action” (CA essay 6), where I first worked it out, and in “Acting for a Reason” (CA essay 7).

<sup>13</sup> John Stuart Mill, *Utilitarianism*, p. 2.

<sup>14</sup> The term is used and the idea explained by Robert Nozick in *Anarchy, State, and Utopia*, pp. 28–33.

<sup>15</sup> This perplexity is expressed, for example, by Samuel Scheffler in *The Rejection of Consequentialism*, p. 82, and by Thomas Nagel in *The View from Nowhere*, p. 178, though neither of them ultimately endorses the idea that deontological restrictions are paradoxical.

<sup>16</sup> Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, chapter 1, pp. 125–31.

philosophers have been prepared to try to pick up that burden, by showing that there are deontological constraints on the use of our productive capacity, hardly anyone has thought to challenge this assessment of where the burden of proof really lies.

### 1.2.2

But it has not always seemed obvious to philosophers that action is production. In Book 6 of the *Nicomachean Ethics*, when trying to work out what art or craft is, Aristotle says:

Among things that can be otherwise are included both things made and things done; making and acting are different . . . ; so that the reasoned state of capacity to act is different from the reasoned state of capacity to make . . . Making and acting being different, art must be a matter of making, not of acting. (NE 6.4 1140a1–16)

According to Aristotle, action and production are two different things. And in the following section, Aristotle remarks on one of the most important differences between the two, namely that:

while making has an end other than itself, action cannot; for good action itself is its end. (NE 6.5 1140b6–7)

Actions, or at least good actions, are chosen for their own sake, not for something they produce.

### 1.2.3

Actually, this is one of three different things Aristotle tells us about why good actions are done by virtuous agents. First of all, in at least some cases an act is done for some specific purpose or end. For instance, Aristotle tells us that the courageous person who dies in battle lays down his life for the sake of his country or for his friends (NE 9.8 1169a17–31). In the same way, it seems natural to say that the liberal person who makes a donation aims to help somebody out (NE 4.1 1120b3), the magnificent person who puts on a play aims to give the city a treat (NE 4.2 1122b23), the magnanimous man aims to reap honors (NE 4.3 1123b20–21), the ready-witted man aims to amuse his audience in a tactful way (NE 4.8 1128a24–27), and so on. At the same time, as I've just mentioned, Aristotle says that virtuous actions are done for their own sake. And finally, Aristotle also tells us that virtuous actions are done for the sake of the noble (e.g. NE 3.7 1115b12; 3.8 1116b3; 3.9 1117b9, 1117b13–14; 3.11 1119b15; 4.1 1120a23; 4.2 1122b6–7).

On an oversimplified conception of moral psychology these will look like three inconsistent accounts of the purpose or aim of virtuous action. But a



little reflection will show why there is no inconsistency here, and at the same time, will throw light on Aristotle's conception of action. What corresponds in Aristotle's theory to the description of an action is what he calls a *logos*—as I will render it, a principle. A good action is one that embodies the *orthos logos* or right principle—it is done at the right time, in the right way, to the right object, and—importantly for my purposes—with the right aim. To cite one of many such passages, Aristotle says:

anyone can get angry—that is easy—or give or spend money; but to do this to the right person, to the right extent, at the right time, with the right aim, and in the right way, *that* is not for everyone, nor is it easy; that is why goodness is both rare and laudable and noble. (NE 2.9 1109a26–29)

The key to understanding Aristotle's view is that the *aim* is included in the description of the action, and that it is the action as a whole, *including the aim*, that the agent chooses.

Let us say that our agent is a citizen-soldier, who chooses to sacrifice his life for the sake of a victory for his *polis* or city. The Greeks seem to think that is usually a good *aim*. Let's also assume that our soldier sacrifices himself at the right time—not before it is necessary, perhaps, or when something especially good, say cutting off the enemy's access to reinforcements, may be achieved by it. And he does it in the right way, efficiently and unflinchingly, perhaps even with style, and so on. Then he has done something courageous, a good action. Why has he done it? His *purpose* is to secure a victory for his city. But the object of his choice is the whole action—sacrificing his life in a certain way at a certain time in order to secure a victory for the city. He chooses this whole package, that is, to-do-this-act-for-the-sake-of-this-end—he chooses *that*, the whole package, as a thing worth doing for its own sake, and without any further end. “Noble” describes the kind of value that the whole package has, the value that he sees in it when he chooses it.

#### 1.2.4

Now this means that Aristotle's view of the nature of action is precisely the same as Kant's. Kant thinks that an action is described by a maxim, and a maxim is also normally of the “to-do-this-act-for-the-sake-of-this-end” structure. Kant is not always careful in the way he formulates the maxims of actions, and that fact can obscure the present point, but on the best reading of the categorical imperative test, the maxim of an action which is tested by it includes both the act done and the end for the sake of which that act is done. It *has* to include both, because the question raised by the categorical imperative test is whether there could be a universal policy of pursuing *this sort of* end