

Library of Ethics and Applied Philosophy 33

Stephen Petro

Rationality, Virtue, and Liberation

A Post-Dialectical Theory of Value



Springer

Rationality, Virtue, and Liberation

LIBRARY OF ETHICS AND APPLIED PHILOSOPHY

VOLUME 33

Editor in Chief

Marcus Düwell, *Utrecht University, Utrecht, NL*

Editorial Board

Deryck Beyleveld, *Durham University, Durham, U.K.*

David Copp, *University of Florida, USA*

Nancy Fraser, *New School for Social Research, New York, USA*

Martin van Hees, *Groningen University, Netherlands*

Thomas Hill, *University of North Carolina, Chapel Hill, USA*

Samuel Kerstein, *University of Maryland, College Park, USA*

Will Kymlicka, *Queens University, Ontario, Canada*

Philippe Van Parijs, *Louvaine-la-Neuve (Belgium) en Harvard, USA*

Qui Renzong, *Chinese Academy of Social Sciences, China*

Peter Schaber, *Ethikzentrum, University of Zürich, Switzerland*

Thomas Schmidt, *Humboldt University, Berlin, Germany*

For further volumes:

<http://www.springer.com/series/6230>

Stephen Petro

Rationality, Virtue, and Liberation

A Post-Dialectical Theory of Value

 Springer

Stephen Petro
Columbia University
New York, New York
USA

ISSN 1387-6678 ISSN 2215-0323 (electronic)
ISBN 978-3-319-02284-0 ISBN 978-3-319-02285-7 (eBook)
DOI 10.1007/978-3-319-02285-7
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013954781

© Springer International Publishing Switzerland 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To My Mother, Diana Borgia-Petro, who
has taught me the meaning of striving,
And to all my close friends, who daily
teach me the value of freedom*

Acknowledgements

This book is the product of a life thus far lived in the relentless pursuit of answers to questions that, whether or not I would have liked them to be posed to me, have been so posed via an infiltrative process initiated and sustained by personal experience, adversity, and the ultimate triumph over adversity. Often viewed as a pursuit of the lofty academic, philosophy, and in particular value theory, is on the contrary a force that constantly beckons every individual at every moment. Questions of value are constantly pushed in our faces; it is up to us whether, upon glimpsing these cognitive disturbances, we take the necessary step back from automated living and resolve to change our lives, that is, to live through a process of coherent intention. For me, this project is of course ongoing, but I owe much gratitude to particular people who helped catalyze and encourage the spirit of inquiry in me that gives my life meaning. I also owe much gratitude to those who continue to sustain and support me, whether actively or symbolically. Some of those I list here will understand precisely why they are on this list, while others might be surprised. To those who are surprised, you might tell me you had nothing to do with any of my achievements (as some of you already have) and that they were all my own doing, but you would be wrong. Had you not been present, I might not have conjured enough strength even to begin to value my own thoughts, ideas, and research topics, let alone discuss or write about them. You are all part of my achievements: stand up and accept your recognition.

First and foremost, I could not have gotten to this point, nor anywhere near it, without the pure and unconquerable love of my mother. She is an indelible part of me and my identity; without her, I would be nothing. Despite her constant travails through rocky waters of uncertainty and suffering, she has poured herself tirelessly into building my strength, my self-confidence, and my sense of intellectual worth against a social milieu that has been largely unsupportive and indeed often adversarial; what an arduous task for a person herself so mercilessly beaten down by the system. Thank you, mother, for making the commitment both to me and my sister to break the cycle. You are beautiful and will remain my best friend forever.

I also could not have arrived at this point without the ardent and steadfast support of my professors at Drew University. In particular, I would like to thank Thomas

Magnell, who has been a thoroughly supportive mentor and many of whose arguments on value theory I have carried over into this work. I would next like to thank Sharon Sundue, who oversaw my undergraduate thesis on the causes of the failure of the antebellum penitentiary system¹ and who has served also as a true intellectual companion. Kant scholar Seung-Kee Lee has also been very influential to the development of my own thought. Indeed, he has acted both as an intellectual mentor and a treasured friend. Darrell Cole also deserves special acknowledgement for his role in overseeing the development of the work in its initial stages; his encouragement, critique, and overall feedback has been invaluable. Jonathan Rose, Lillie Edwards, Edward Baring, James Carter, and John Lenz deserve special recognition for the intellectual challenges, in general, they have repeatedly posed to me and for the intellectual companionship they have offered. These individuals have demonstrated what a more egalitarian and less hierarchical student-professor relationship is capable of producing. Finally, I must thank with sincerest gratitude Anthony Morinelli of the Shipley School, for initially introducing me to the formal study of philosophy and ethics; it was he who provided me with the first step on my thousand-mile journey.

I must also acknowledge individuals who have generously supported the financing of my education. Special thanks to Laura and Karen Borgia, my sister Barbara Petro Escobar, Greg Kleiber, Esq., Dr. Mort and Mrs. Susan Paterson, Mr. Charles and Mrs. Barbara O'Reilly, Mr. Michael and Mrs. Allyson Lunny, Mr. Rob Anderson, Mr. Girish and Dr. Sapna Havildar, and the organization Social Grad. Your support has played an important role in making the production of this book possible.

¹ 2012. *Broken Men: Poor White Manhood and the Failure of the Antebellum Penitentiary System*. Madison: Drew University.

Contents

1	Introduction	1
2	Rethinking Rationality	9
2.1	The Reconciliation of Ethical Rationalism, Ethical Naturalism, Virtue Ethics, and the Biological and Social Sciences	9
2.1.1	Reasserting the Compatibility of the Methodologies of Value Theory and of the Sciences	10
2.1.2	Reasserting the Compatibility of Virtue Ethics and Deontological Rationalism	24
2.2	The Failure of Axiological Anti-foundationalism	29
2.2.1	The Ideological Dangers of Anti-foundationalism	29
2.2.2	The Problem of Indeterminacy: Amartya Sen’s Theory of Justice	32
2.2.3	Karl-Otto Apel’s Rebuttal to the Münchhausen Trilemma and Its Relevance to Value Theory	35
2.3	The Concept of Rationality: Toward a Universal Model	36
2.3.1	Concepts, Conceptions, and the Possibility of a Universal Model	37
2.3.2	“Humaniqueness,” Social Cognitive Theory, and a Neuroscientific Account of Judgment	40
2.3.3	Inadequacies and Limitations of Competing Accounts	45
2.3.4	The Necessary Methodological Preconditions of Universal Applicability	53
2.3.5	The Model Explicated and Analyzed	54
2.3.6	Consequences for a Doctrine of Liberation	61
2.4	Concluding Remarks	63

3	Rationality and Dialectical Necessity	65
3.1	Prescription, Preference, and Dialectical Contingence	65
3.1.1	A Refutation of Harean Anti-descriptivism	67
3.1.2	Hare’s Dialectical Method	79
3.2	Developing a Method of Justification	91
3.2.1	The Problem of a Justificatory Method	91
3.2.2	Problems in Walton’s Model of Justification	93
3.2.3	The Justificatory Model Explicated and Analyzed	97
3.3	A Sound Positive Account, Part I: An Analysis of Gewirth’s Ethical Rationalism	106
3.3.1	The Basic Elements of Gewirthian Theory	106
3.3.2	The Premises and Conclusion Reconstructed and Analyzed	111
3.3.3	Applications to Animal Ethics and the Principle of Proportionality	114
3.3.4	Applications to Positive Rights	120
3.3.5	Applications to Classic Moral Dilemmas, the Principle of Double Effect, and Intention	122
3.3.6	Position A: Incompleteness and Equal Justifiability	127
3.3.7	Position B: The Self-Contract	129
3.3.8	Position C: Double Effect, Contrary Double Effect, and Synthetic Double Effect Derived from the PGC	133
3.3.9	Conclusion and Synthesis	135
3.4	A Sound Positive Account, Part II: An Analysis of Habermas’s Discourse Ethics	139
3.4.1	Critical Theory, Interaction, and Communicative Ethics	142
3.4.2	Habermas, Gewirth, and the Inescapability from the Monological	147
3.4.3	Clarifying the Mirror Image	150
3.4.4	Epistemological Uncertainty and Alter-Contradiction: A Reconstruction of the Habermasean Argument	153
3.4.5	Applications to Moral Dilemmas: The Right to Discourse	155
3.4.6	Applications to Property Theory	157
3.4.7	The Baker and the Starving Man	162
3.4.8	The Right to Discourse, Self-Sacrifice, and the Antique Clock	165
3.4.9	Historical Arguments for a Reinterpretation of Habermasean Ethics	169
3.4.10	Concluding Remarks: The Single-Person Problem	172

- 4 The Dialectical Structure of Value Judgments** 175
 - 4.1 The Dialectical Structure of “Ought” and “Must” 175
 - 4.1.1 The Harean Account of Imperative, Indicative,
and Conditional Statements 175
 - 4.1.2 Actuality, Subjunctivity, and Conditionality 178
 - 4.1.3 Implications of Gewirthian Theory
for Deontic Meaning 188
 - 4.2 The Dialectical Structure of Rights and Duties 190
 - 4.2.1 Relativism and the Correlativity of Rights,
Duties, and “Must” 190
 - 4.2.2 Rights, Duties, and Anachronism 193
 - 4.2.3 Concluding Remarks 194
- 5 Virtue and the Search for Intrinsic Goodness** 197
 - 5.1 Magnell’s Challenge 197
 - 5.1.1 Probing the Predicative-Attributive Distinction 198
 - 5.1.2 The Definitive Criteria for Value Intrinsicity 200
 - 5.2 Problems in Searle’s Epistemology of Function 202
 - 5.2.1 The Social Construction of Function 202
 - 5.2.2 The Problem of Classification in the
Absence of Essence 205
 - 5.2.3 Functions, Mental States, and the Background 209
 - 5.2.4 Concluding Remarks 210
 - 5.3 The Life Framework: The Significance of Foot’s
Virtue Theory 212
 - 5.3.1 Function, Necessity, and the Grammar of Goodness 213
 - 5.3.2 The Magnitude of the Implications 216
- 6 Beyond Dialectical Necessity: Assertoric Necessity
and the Grammar of Goodness** 221
 - 6.1 Reflexive Intrinsicity and the Teleologically Comparative
Tendential Necessity of Functions 222
 - 6.1.1 Assaying Some Peculiarities 222
 - 6.1.2 The Good-for Thesis 223
 - 6.1.3 The Reflexive Intrinsicity of Goodness 224
 - 6.1.4 Tendency, Necessity, and Function 226
 - 6.1.5 Applications to Goodness 232
 - 6.1.6 A Naturalistic Definition of Goodness 234
 - 6.1.7 First Proof of Intrinsic Goodness 236
 - 6.1.8 Second Proof of Intrinsic Goodness 238
 - 6.1.9 The Universality of Goodness and the Comparison
of Goods 240
 - 6.1.10 Objection to Betterness and Reply 242

- 6.1.11 A Naturalistic Definition of Betterness 243
- 6.1.12 Objections and Replies 246
- 6.1.13 Teleological Comparativeness 248
- 6.1.14 Objections and a Reply via the Argument
to the Commensurability of *Teloi* 250
- 6.1.15 Comparison and the *Telos* of Betterness 252
- 6.1.16 Another Objection to the Commensurability
of *Teloi* and a Reply 256
- 6.1.17 Conceptual Connections to the Non-Descriptivist
Fallacy 259
- 6.1.18 Objections on Grounds of Inherent Relativity
and a Reply 260
- 6.1.19 Objection on Grounds of Failure to Prove that Value
Is a Natural Property and a Reply 262
- 6.1.20 Objection on Grounds of Dialectical
Necessity and a Reply 263
- 6.1.21 Objection on Grounds of the Intrinsicity-Finality
Distinction and a Reply 264
- 6.1.22 Final Considerations: The Inadequacy of Summation
and Aggregative Models of Value 265
- 6.2 The *Summum Bonum* 267
 - 6.2.1 From Value to Rationality: The Fundamental
Link Elucidated 268
 - 6.2.2 Justifying a *Summum Bonum* 272
 - 6.2.3 Logical Judgment 275
 - 6.2.4 Conceptual Synthesis 276
 - 6.2.5 Conceptual Abstraction 278
 - 6.2.6 Freedom 279
 - 6.2.7 Wisdom and the Unity of the Virtues 281
 - 6.2.8 Objections and Reply 282
 - 6.2.9 The Question of Attainability 284
 - 6.2.10 *Phronesis, Eudaimonia, and Nirvana* 287
 - 6.2.11 The Question of Individual *Telos* 290
 - 6.2.12 Some Implications for Aesthetics 291
 - 6.2.13 The Question of Individual *Telos* Revisited:
The Relevance of Phenomenology 292
 - 6.2.14 Gewirth and Habermas Revisited 296
 - 6.2.15 The Single-Person Problem Resolved 298
 - 6.2.16 The Final Assessment: Choosing a Model 299
- 6.3 Conclusion 300
 - 6.3.1 The Eternal Nature of the *Summum Bonum* 300
 - 6.3.2 Enlightenment and Liberation 302

7 Conclusion 309

 7.1 The Eternal Nature of the *Summum Bonum* 309

 7.2 Enlightenment and Liberation 310

 7.3 Preliminary Implications for Political Theory 313

Bibliography 319

Chapter 1

Introduction

Some of the oldest questions in the history of the world are questions of value: “Who am I?,” “What should I do?,” and “What kind of person should I be?” Although these questions have persisted for so long and thus seem to some endless and pointless, they are indeed still well worth asking. In the first instance, that these questions have persisted for so long merely attests to their complexity and depth. In the second instance, such a claim of pointlessness also presumes that there has been nothing we could call progress in answering these questions. Indeed, there has been much progress, and in fact there seems to be something awry with the claim that asking the question “What is the point?” (a question of value) is itself pointless or valueless. There is a function or *ergon* revealed within this utterance which seems in itself to elucidate and elicit value. One of the goals of this book is to explain precisely why this is the case.

In the history of ethical thought, we can delineate some major trends. In the ancient world, as Alasdair MacIntyre and others rightly claim, the idea of virtue was paramount. The question “What kind of person should I be?” set the ethical paradigms. From Plato, Aristotle, Epicurus, Diogenes of Sinope, and Cicero to such thinkers as Lao Tzu, Confucius, Vasubandhu, and the Buddha, questions of personal identity and personal fulfillment shaped ethical questions. It is, however, limiting to state that these were merely *ethical* questions. In fact, the question “What kind of person should I be?” did and does encompass far more than what moderns would consider ethics or morality. The kind of person I should be certainly depends on *who I am essentially* and what possibilities there are for me. When I ask this question, therefore, I ask more than what kind of member of society or moral person I should be; I inquire, rather, into my very soul. “Who am I?”: Shall I seek to be a knowledgeable and wise person, and if so, how shall I do it? Shall I seek to be a person of healthy body? Shall I seek to be a compassionate, courageous, and prudent person? Shall I seek to be a loving person, a person who does not live alone but with true friends and deep companionship? Shall I seek to be a content person, a person full of mental peace? If the answer to all of this is “yes,” then I must conclude that being a good person, the best person *I* can possibly be, includes more than simply how I treat other people. It includes, in addition, a profound and

earnest engagement in a process of self-analysis and self-discovery, ultimately leading to the determination of who and what *I* am, a process which is indispensable for answering the aforementioned questions. We must also acknowledge, however, that these questions are inextricably tied to how we treat others and, importantly, how we *must* treat them.

The virtue paradigm in Western thought persisted to a large extent throughout the medieval period but began to fade in earnest with the Enlightenment. As many virtue theorists have acknowledged, this was surely a tragedy. It was not, however, a complete tragedy for ethics. As the concepts of duty and right became prominent and superseded classical notions of virtue, most saliently in the work of Kant, Western thought lost sight of most of these aforementioned questions and focused almost exclusively on the far more limited question “What should I do?” With this hegemony came a new paradigm for ethics, and with this new paradigm came the abnegation of the self. The main questions became “What are my duties?” and “What are my rights?,” and from this emerged a one-dimensional conception of goodness, personhood, and liberty. The simplistic idea arose in political theory that I may do anything I *want* so long as I do not infringe upon others’ rights. Because there was a vacuum of space left in the wake of the disappearance of the other questions, personal flourishing fell back on wants, desires, and preferences, while such philosophers as Hobbes and Mill claimed that the only way one could even conceive of value’s reality was in terms of desire, preference, or pleasure. The stage was thus set for the worship of *pleonexia*. The ideas of duty and right, however, were implicitly tied to the original questions of ancient thought, namely through the concepts of *ought*, *should*, and *must*. A sole emphasis on deontic concepts hid from view the necessities to which they pointed and upon which they relied for their meaningfulness: those of value and virtue. While virtue theorists often denounce the concepts of rights and duties, therefore, they overlook, as I will show, that these concepts ultimately derive from a system of virtue and excellence. The total rejection of the Enlightenment or of deontic concepts, therefore, is a mistake.

In the twentieth century, the stage was set for a slew of different strains of thought. While emotivism and various forms of anti-realism persisted early in the century, virtue once again peaked through the mist via analytic philosophy. The work of Philippa Foot, G. E. M. Anscombe, Peter Geach, John McDowell, MacIntyre, and others provided a new foundation for the ancient questions of which humanity could no longer be deprived. Most of them, however, made the mistake of rejecting the advances of deontology and rationalism that had taken hold during the Enlightenment. Alongside this aretaic trend came other philosophers who, in the spirit of Kant, attempted to provide and did successfully provide sound bases for notions of rights, duties, and justice. The most successful of these theories have been dialectical, or centered on the statements that individual agents make. R. M. Hare, Alan Gewirth, and Jürgen Habermas are by far the most prominent within this tradition of dialectical justification. There are three approaches to dialectical justification which each philosopher represents. The two main categories are dialectical contingency and dialectical necessity, of which Hare represents the first and Gewirth and Habermas represent the second. Dialectical contingency is a

methodological approach which evaluates the truth values, or at least the validity, of normative propositions via the normative statements or claims that agents incidentally make (e.g.: “I prefer living to dying”), while dialectical necessity bases itself upon the proposition that the truth of normative statements is founded upon statements or claims which agents must make insofar as they are agents (e.g.: “I make choices.”); from a pragmatic standpoint, they are inescapable. Within the category of dialectical necessity are the monological approach and the dialogical approach, which Gewirth and Habermas represent respectively. The monological approach is founded upon the proposition that legitimate normativity is constituted by those statements or claims which agents make without any interaction or intervention implied on the part of other agents (e.g.: “I make choice x .”). This is contrasted with the dialogical approach, in which the truth values, or at least the validity, of normative statements rely upon particular interactions within a discursive milieu (e.g.: “I make choice x but cannot determine its legitimacy without your evaluation, assent, and possibly consent.”). The consideration of the relative merits and insights of such dialectical theories will constitute a main analytical goal of this book. As we shall see, while powerful and widely applicable, the dialectical model alone, while capable of grounding a sound theory of ethics, is insufficient to ground a sound theory of *value*. Something more, namely a recognition that the aforementioned questions are inherently aretaic and can therefore be answered only via an aretaic elucidation of the nature of value, is required.

After Gewirth wrote his main work, *Reason and Morality*, numerous philosophers took note of it, many in support of his theory and many in rejection of it. Many of these objections, criticisms, and suggestions, as well as Gewirth’s replies to them, later appeared in Edward Regis’s anthology, *Gewirth’s Ethical Rationalism*. The discourse, however, did not stop there. Deryck Beyleveld shortly followed with his work, *The Dialectical Necessity of Morality*, in which he not only defended Gewirth against further criticisms not published in *Gewirth’s Ethical Rationalism* but also laid out, in rich and painstaking detail, the logical structure of Gewirth’s theory.¹ Beyleveld has rightly noted that it seems, at least superficially, “that it is a case of Gewirth versus the rest of the philosophical world.” As the last decade or so has seen a general silence, with a few exceptions, on Gewirth’s ethical theory, it is well worth taking up Beyleveld’s advice that “Gewirth’s supporters. . . stand up and be counted.”² Likewise, I am also strongly in support of Habermas’s communicative ethics, although as I will explain in later chapters, I do not accept the fundamental epistemological basis upon which he claims to justify the theory. In fact, as I will show, Habermas’s and Gewirth’s theories share more in common with each other than they differ; they are what I will call

¹ Deryck Beyleveld, *The Dialectical Necessity of Morality: An Analysis and Defense of Alan Gewirth’s Argument to the Principle of Generic Consistency*. (Chicago: The University of Chicago Press. 1991); Alan Gewirth, *Reason and Morality*. (Chicago: The University of Chicago Press. 1978); Edward Regis, *Gewirth’s Ethical Rationalism: Critical Essays with a Reply by Alan Gewirth*. (Chicago: The University of Chicago Press. 1984).

² Beyleveld, *The Dialectical Necessity of Morality*, 5.

“mirror images” of each other. As such, their combined accounts provide powerful solutions to particular perennial issues in moral philosophy, including those concerning animal rights, property rights, the justification of the double-effect principle, and self-sacrifice for the greater good. I will argue that Hare’s theory, on the other hand, is fundamentally unsound. His theory, although expressly not deontological, is important for another reason: his account helps to demonstrate that certain judgments concerning value, and more specifically ethical value, are false via pragmatic and dialectical inconsistency. The theoretical aspects of Harean Universal Prescriptivism, therefore, aid us in our journey toward justifying a methodology, guided by certain sound criteria, whereby we can determine, *before* determining which kinds of normative judgments are *true*, which kinds of normative judgments are *false*. This step, I believe, is important for narrowing the range of possible sound normative models and thus for justifying our progression toward the particular approach I endorse.

Despite the great successes in the realm of dialectical ethical theory, these theories suffer from various problems, including their failure to address the other important questions I have mentioned. These questions cannot adequately be answered by appeals to rights and duties, or in Habermas’s case, in terms of deontological justice. We must, however, realize that rights and duties give us part of the answer, if not the entire answer. As one of the main themes of this book, I will argue that deontic and aretaic concepts are logically and semantically intertwined and that neither the dialectical deontological theories nor virtue theories thus far have clearly and succinctly elucidated these conceptual links. As I will argue, the questions “What is value?” and “What is goodness?” seem to have had no significant resolution partly due to this oversight. When we closely analyze these conceptual links and discover the common bond that exists between dialectical deontology and virtue theory, we are then able to discover exactly what value and goodness are. Indeed, we are able to *define* goodness. This is a bold claim, and the definition of goodness is a feat that has not been fully tried since the early to mid-twentieth century. This feat, however, can be accomplished.

It is a central purpose of this book to develop a full-fledged value theory. I am not, in other words, interested in discovering merely a sound ethical theory. As I will argue, once we leave pure dialectical deontology behind and relate the vital elements of it to virtue theory, we are then able to form a theory of value more generally, one which adequately answers the aforementioned questions. It is my hope that not only ethics but also such fields which address questions of value as aesthetics, economics, business studies, history, law, and political theory will be able to find grounding within this theory. The theory is post-dialectical in that, although it takes vital concepts from the dialectical theoretical framework, it is ultimately independent of this framework and instead relies on a framework of virtue. It achieves this post-dialectical status via the reapplication of certain dialectical insights to an assertoric model of value judgment and through the demonstration, via an exhaustive and novel analysis of the concepts of goodness and betterness and an appeal to a pseudo-Carnapian conception of the empirical-linguistic or linguistic-conceptual framework, of how such a model is justifiable

independently of the judgment of agents. Integral to the establishment of such an assertoric model of normativity will be my account of rationality. Rationality, I will argue, is not merely a deductive and inductive logical phenomenon. Although these are vital components of rationality, rationality is also importantly experiential and phenomenological. The failure of most value theorists to consider the normative implications of experiential rationality, I will argue, has resulted in modern theoretical models which either offer only partial answers to the aforementioned questions or which ignore them completely. There are, as I will demonstrate, profound normative consequences for a consideration of the aspects of rationality that concern what might be considered our “raw,” non-intellectualized experience, as opposed to our inductive and deductive judgments. These experiential components remove us from the one-dimensional conception that many modern theories, including rational- and social-choice theory, have offered us. The reincorporation of these elements leads us to a view of rationality that is inherently anti-elitist, a far different conception than that which the critical theorists and post-structuralists have also offered us. Rationality, in fact, is so crucial for instantiating value, understanding who we really are, and for finally setting ourselves free. Far from oppressive, rationality leads the path to liberation.

This book is comprised of seven chapters. In Chap. 2, I set out on the journey toward value by addressing issues concerning the idea of rationality. In Sect. 2.1, I challenge some prominent views concerning the supposed incompatibility of ethical rationalism, ethical naturalism, virtue ethics, and the methods of the biological and social sciences. I begin to hint at the common approaches and conceptual connections between these schools of thought which I show in later chapters renders them intrinsically compatible. This demonstration of theoretical compatibility ultimately aids in revealing those concepts that are central to rationality and the ways in which dialectical deontological rationalism most accurately elucidates them. In Sect. 2.2, I address the view of anti-foundationalism prominent among contemporary philosophers. I argue against these claims and contend that a foundationalist framework is necessary for addressing questions of value and rationality. In Sect. 2.3, I set out to characterize rationality. In order to do so, I first challenge various existing views on what rationality is and then set out to establish my own characterization using evidence from the fields of biological anthropology, social cognitive theory, neuroscience, phenomenology, and philosophy more generally. I argue for what I claim is the *concept* of rationality, as opposed to the various *conceptions* of rationality, to which everyone implicitly refers in invoking rationality. As the most general and generic characterization possible, I argue, we can capture the essential elements of the linguistic-conceptual framework that rationality comprises which are, as such, supremely relevant to the establishment of legitimate normativity. This analysis and refiguring of rationality also proves vital, in later chapters, to our understanding of what I will claim are the most relevant and important elements of dialectical ethics. It is ultimately this reconceptualization of rationality that will undergird and legitimate my move from a merely dialectical model of value to an assertoric one which incorporates these vital components of dialectical theory. In Chap. 3, I tackle dialectical

deontology and dialectical utilitarianism directly. In Sect. 3.1, I first assess and critique Hare's universal prescriptivism. I conclude that, although his theory suffers from serious methodological and theoretical flaws, certain ideas and constructs from his framework aid us in an important task that is all too often neglected in ethical theory: the establishment of a negative account of justification. In Sect. 3.2, I show that, by establishing a negative account of justification, we can eliminate certain types of value judgments by demonstrating them to be false. We can also, I claim, demonstrate who has a burden of proof in an ethical argument or an argument concerning value more generally. From here, we can more strongly found positive accounts of ethical justification. In Sect. 3.3, I explore the major merits of Gewirth's theory and demonstrate, through a somewhat different approach than that which is currently propounded by Gewirth's main adherents, that his theory has wide-ranging application not only to prominent moral dilemmas but also to animal ethics and other issues such as the significance of intention. My approach to these issues, I believe, is closest to the Gewirthian dialectical approach because it addresses such issues via direct analyses of the implicit or explicit statements made by agents in such situations and a subsequent determination of which statement among all others entailed by particular acts is consistent with Gewirth's supreme principle of morality, the Principle of Generic Consistency (PGC). This methodological approach, I will argue, is the most meticulous and, as such, leaves less room for error in determining what exactly it is that follows from the PGC. It is partly for these and the aforementioned reasons, I argue, that the dialectical model is so theoretically powerful. In Sect. 3.4, I show that Habermas's theory is also sound in these ways but that his theory is ultimately not dialogical but instead monological. As such, his theory is a mirror image of Gewirth's in that Habermas's begins on a premise of alterity whereas Gewirth's begins on a premise of self-reference. As such, Habermas's theory fills in gaps that Gewirth's theory seems to leave vacant, namely as regards highly nuanced interpersonal conflicts such as those concerning the seemingly ambiguous ownership status of particular items of personal property, as well as the tenuous nature of interpersonal sacrifice. It is what I will call the Right to Discourse that very substantively resolves such dilemmas. I conclude, however, that Gewirth and Habermas's theories are inadequate because fundamental semantic and conceptual gaps remain which only virtue theory, founded upon my reconceptualization of rationality, necessarily fills. In Chap. 4, therefore, I analyze the dialectical structure of value judgments and conclude that the logical and semantic connections between judgments of goodness and deontic judgments are inextricably connected in such a way that judgments concerning rights and duties cannot be separated from judgments concerning goodness, value, and virtue. Whereas deontologists generally claim that virtue theory is dependent upon deontic concepts, however, I show that the reverse is true; the necessity and conditionality upon which deontic judgments rest are ultimately reliant upon values and value systems for their intelligibility. Without this underlying structure of direct value judgment, deontic judgments are completely empty and vapid. In order to elucidate this, in Sect. 4.1, I analyze "ought" and "must" grammatically and logically and, in Sect. 4.2, I tie this analysis

to an analysis of the concepts of rights and duties. In Chap. 5, the analysis builds upon the last three chapters and culminates in an account of the nature of goodness and betterness themselves. I utilize these previous analyses, incorporating fundamental deontic and aretaic concepts as well as fundamental insights from dialectical theory, to demonstrate that goodness and betterness are what I call “reflexively intrinsic” and that, by extension, questions concerning what is good or valuable are not by any means open ones, as many anti-realists and non-cognitivists contend. In Chap. 6, I conclude that rationality is fundamental to answering the questions “What is goodness?,” “What is value?,” “What is good?,” and “What is best?” Building on the work of Foot and Thomas Magnell, I explain that rationality constitutes a conceptual framework which makes such questions possible, intelligible, and meaningful. As such, the very concept of value has implicit reference, albeit in a very particular manner, to rationality via an assertorically necessary grammar of value. The concepts contained within rationality, among them *phronesis*, conceptual abstraction, conceptual synthesis, and freedom, provide a framework that makes questions of value, commonly portrayed as an endless chain of indeterminate and unanswerable prompts, ultimately subject to a final arbiter. As it turns out, this chain of questions has a final link, and that final link is, in a very non-trivial sense, the assertoric (and not merely dialectical) necessity of value itself.

Chapter 2

Rethinking Rationality

In this chapter, I attempt to establish the preliminary foundations which demonstrate the possibility of a post-dialectical, assertoric value theory. I set out to demonstrate that the schools of ethical rationalism, ethical naturalism, virtue ethics, and the methods of the biological and social sciences, while often opposed theoretically, in fact offer each other support and justification. In order to set this relationship right, and before exploring dialectical theory itself, I argue that it is necessary to radically revise the concept of rationality to one that has joint scientific, social-scientific, and phenomenological support. Classic analytic and critical-theoretical characterizations, I argue, are inadequate.

2.1 The Reconciliation of Ethical Rationalism, Ethical Naturalism, Virtue Ethics, and the Biological and Social Sciences

In this section, I offer a brief reinterpretation of the relationship between ethical rationalism, ethical naturalism, virtue ethics, and the biological and social sciences. A reconciliation of these three schools of thought sets the foundations not only for my exposition and analysis of dialectical ethical theory but for my account of rationality, which will undergird my establishment of post-dialectical value theory.

Since the nineteenth century, and even to some extent prior, ethical rationalism has been subject to intense scrutiny and criticism. Although the ethical rationalist paradigm has benefited greatly from this and has improved on several fronts, there are some principal barriers which have severely hindered it. In the first instance, ethical rationalism currently contends with what might be called the growing “turn to virtue” in ethical theory, much of which stands in opposition to rationalistic conceptions of ethics. In the second, ethical naturalism increasingly often seems to stand aloof with respect to both ethical rationalism and virtue ethics, being seen perhaps as an unnecessary discussion to broach in pursuit of a sound ethical theory,

perhaps for fear of rehashing bothersome metaphysical baggage. Yet still, ethics in general has too often been misrepresented, mischaracterized, or implicitly ridiculed in the sciences, leading to a gradual withdrawal of ethical naturalists, particularly among scientific integrationists and virtue theorists, from the province of ethical rationalism and leading some, disillusioned by what they apparently perceive as unscientific (and therefore presumably irrational) methods, to veer off in the direction of an inductively derived account of normative theory. This has been especially unfortunate, since ethical rationalism and certain forms of ethical naturalism complement each other in vital ways. In order to address meta-ethical theory in a meaningful way, it is important to revisit these respective paradigms of naturalism, rationalism, and the biological and social sciences and to call them into collaboration. In order to establish the basis for this collaboration, it will be important, first, to confront some competing paradigms.

2.1.1 Reasserting the Compatibility of the Methodologies of Value Theory and of the Sciences

With regards to the theoretical challenges to ethical rationalism, there are a particular few which demand our attention. Although ethical rationalism need not, itself, be founded on cognitivism or descriptivism, there are some non-cognitive and non-descriptivist frameworks, such as emotivism and intuitionism, which, when positing a moral standard, often emphasize altruism and empathy as primary motivators and justifiers of action. Increasingly, these frameworks have drawn upon evidence provided by the fields of biology and psychology, and this has been primarily for the purpose of clarifying the nature of sentiments associated with moral action. Such work, while fruitful in elucidating the nature of empathy and altruism, falls short in many regards. By appealing to biological and psychological processes, some emotivist and intuitionist theories fall prey to the temptation to assert relationships between these natural phenomena and goodness, simultaneously crossing the line between descriptivism and non-descriptivism and failing the open-question test, at least where it applies. Such frameworks are sometimes supported by theories asserting the existence of a universal moral grammar or, otherwise, universal moral capacities. While the current evidence might support such hypotheses, ethical theories that use such evidence as a justifications for specific actions or, otherwise, to bridge the supposed is-ought gap, seem to fall prey to at least some version of the open-question test and to lack any substantive framework of formal justification. As mere observations or explanations of phenomena, they rarely touch upon the substance of ethical concepts as such, namely through analyses of ethical meaning, syntax, or phenomenology. Especially seeing as though most of these normative undertakings lead also to some form of emotivism or intuitionism, they leave so many questions unanswered about the nature of value that it is, in the final estimation, difficult to conceive of these as explanatorily powerful in any sense. Such

emotivist and intuitionist theories, therefore, unless significantly clarified and improved in the future, seem to lack fundamental justifiability as *normative* theories of value.

To be fair, biological anthropologist, Marc Hauser, who originated the hypothesis of universal moral grammar, does not hold a purely emotivist or intuitionist position.¹ There are, however, certain features of his method of argumentation and justification, the most salient being his inductive method, that are common to other scientifically based meta-ethical theories. Such ethical naturalists as William Rottschaeffer have also endorsed the use of inductive methods for attaining moral knowledge. While the integrationist approach is a sensible one, both Hauser and Rottschaeffer focus their attention heavily toward certain aspects of moral agency, such as empathy, altruism, and base- and behavioral-level judgment, without fully explaining their relevance to *justifiable* moral judgment. This approach, coupled with an off-hand rejection of deductive methods, brings them into the difficult position of attempting to justify moral assertions via an inductively derived pluralism. Predictably, this leaves their respective accounts with much to be desired in the way of firm theoretical foundations, namely sound formal and coherent contentual frameworks. Even Nancy Snow, whose social-psychological analysis and vindication of virtue has been groundbreaking for virtue ethics, suffers from shortsightedness in attempting to establish virtue theory on the empirical phenomenon of social intelligence. Once again, the conception of virtue she sets forth is normatively presumptuous and merely makes descriptive claims via an inherently question-begging inductive argument.² This is not to say that empirical methods are irrelevant to meta-ethics; it is simply that the methodology that integrationists have thus far employed omits certain important *a priori* and phenomenological elements. As such, such an empirical methodology cannot constitute the grounds for value justification and legitimacy.

In the eyes of some theorists, however, this may seem a naïve course of argumentation, for it seemingly assumes that there is no possible method of empirical justification that could yield a sound account of normative ethics. In effect, it omits the possibility that ethical justification just is a form of empirical justification and seemingly assumes that deductive methods found all truth claims. More sophisticated integrationist arguments have implicitly attacked just these supposed assumptions. In the field of evolutionary ethics, for instance, Robert J. Richards has offered an unique and compelling argument invoking a loose idea of universal moral grammar and citing evolutionary theory to support his claims. Although not explicit about his reservations concerning *a priori* methods of justification (especially seeing as though he makes quite the homage

¹To be clear, Hauser founded the theory of universal moral *grammar*, apparently modeled after Chomsky's linguistic theory; he is by no means, however, the first to set forth a biological-anthropological or evolutionary account of universal moral sentiments or values as per value nativism.

²Nancy E. Snow, *Virtue as Social Intelligence: An Empirically Grounded Theory*. (New York: Routledge, 2010).

to Kant), he clearly questions the necessity of such methods for the attainment of normative truth. As he states:

Frameworks, their inference rules, and their principles are usually justified in terms of intuitively clear cases—that is, in terms of matters of fact. Such justifying arguments, then, proceed from what people as a matter of fact believe to conclusions about what principles would yield these matters of fact. This method of justifying norms is not confined to ethics. It is quite commonly used in all normative disciplines. In aesthetics we justify principles of artistic value by showing that they would yield the conclusion that the Madonna of Leonardo is quite beautiful, and not really to be compared with the Madonna of MTV.³

Essentially, then, Richards's is an argument concerning the nature of normative judgment, the central claim of which is that normativity just is a conglomeration of systematic judgments about empirical facts. Indeed, as a few wayward philosophers of mathematics have done concerning the existence of numerical entities,⁴ Richards takes this argument all the way back to logic itself. As he states: "In logic, this same strategy has established *modus ponens* as the chief principle of the modern discipline: *modus ponens*. . . renders the same arguments valid that rational men consider valid. But this strategy for justifying norms utilizes empirical evidence, albeit of a very general sort."⁵ Based upon an evaluation of the rest of his essay, I believe it is a fair interpretation of this claim to state that Richards is here asserting that logic itself is empirically situated and that the *a priori* is thus, most problematically, in some way contingent. It is through this contextual understanding of his conceptualization of logic and the *a priori*, therefore, that we must interpret his claims. "Quite simply," he states, "the strategy recognizes what William James liked to pound home: that no system can validate its own first principles. The first principles of an ethical system can be justified only by appeal to another kind of discourse, an appeal in which factual evidence about common sentiments and beliefs is adduced."

On the grounds of his premises concerning the empirical status of logic and justification, he develops the claim, interestingly but counter-intuitively, that the moral realm is constituted by empirical and factual evolutionary contexts that are primarily grammatically situated. A moral assertion, he claims, is no different than any ordinary scientific claim. He argues, "Just as the context of physical nature allows us to argue 'Since carbon dioxide has built up, atmospheric temperature ought to increase,' so the structured context of human evolution allows us to argue 'Since each person has evolved to promote the community good, each ought to act altruistically.' And here I've gone from a factual premise about evolution to an ought proposition, without, I believe, any fallacy. The rule that allows me to join the premise with the conclusion is one that governs the usage of the term 'ought.'" Anticipating likely objections, Richards asks "Is the word 'ought' used any differently here than in the proposition

³ Robert J. Richards, "Birth, Death, and Resurrection of Evolutionary Ethics." *Evolutionary Ethics*. Ed. Matthew H. Nitecki and Doris V. Nitecki. (Albany: State University of New York Press, 1993): 127.

⁴ I am thinking primarily of Mill and Quine.

⁵ *Ibid.*

about the greenhouse effect?” His answer is quite intriguing indeed: “The answer is yes and no. The ‘ought’ of the greenhouse example is not a moral ought. What makes the conclusion a moral-ought conclusion is that the structured context from which it is derived is that of the evolution of altruism. The ‘ought’ derived from the structured context of human evolutionary formation, then, will be a moral ought precisely because the activities of promoting the community good and of approving altruistic behavior constitute what we mean by being moral.”⁶ This is certainly an interesting and potentially very important claim, but my primary objection to Richards lies in his dubious account of *a priori* judgment. If the mere brute facts of evolutionary theory are so fundamentally related to our normative judgment, then why argue for a universal moral *grammar* at all? Put a different way, it seems odd for Richards to argue for a normativity founded upon the empirical facts relating to evolutionary theory but simultaneously to argue for a theory of ethical meaning via the idea of contextualization and conceptual frameworks. Richards may wish to disavow deductive methods, but it seems he will have to accept them in some form if he is to accept his own theory, in which the idea of meaning seemingly stands dominant to biology. Richards will likely object that, as he stated beforehand, analyticity is merely an empirical phenomenon. Of course, he takes as proof of this the supposed truth that all principles are contextual and can never be self-justifying. This claim, however, is a blatant paradox, at least so long as we are to take this claim as indubitably true. Indeed, if we are to take this claim as true at all, even within the loosest possible framework we can imagine (say, that of the conceptual or empirical itself), then we are committed to the conclusion that justification, indeed this very ultimate justification, is possible independently of empirical contingencies, for this claim would hold independently of the existence of sapient beings and even in the most fundamental, all-encompassing conceptual frameworks. Thus, my rebuttal to Richards’s admittedly more sophisticated integrationist argument turns out to be a half-hearted one, for, while I am in tentative support of his project for a universal moral grammar, his claims ultimately extend further than this into the realm of universal moral meaning, which, whether framework-dependent or not, leads implicitly to a claim of analyticity concerning such universal moral judgment. This part of his project, which I do support, fundamentally conflicts with his claim to the evolutionary and thus empirical foundations of normative justification. While some integrationists do veritably abandon aprioristic methods in favor of aposterioristic ones, therefore, it can be seen that some, such as Richards, attempt to have their cake and eat it too. While Richards’s account might ultimately have great promise for value theory, especially as regards the important ideas of conceptual, linguistic, and empirical frameworks, his account ultimately has very little to do with evolutionary theory, and evolutionary theory likewise has little to do with his account, at least directly.

Gewirth, in fact, points out this very incoherence inherent in attempts to found normative theory in evolutionary biology. In his aptly named article, “How Ethical is Evolutionary Ethics?,” Gewirth places into question whether what scientific accounts

⁶ *Ibid.*, 129.

of normativity are attempting to examine is normativity or ethics at all. He posits what he calls a “discontinuity” between the claims of evolutionary ethicists concerning the biological foundations of values on the one hand and the cognitive-perceptual basis of values on the other. His arguments directly address the dilemma that I have claimed is inherent in the normative models of Richards, Rottschaeffer, Hauser, and other such integrationists. As he states: “Evolutionary ethics, being a causal theory, cannot account for the voluntariness and intentionality of moral ‘oughts.’ There is a discontinuity between the intentionality of such ‘oughts’ and the causal necessity that the evolutionary theory attributes to what it depicts as ethical behavior. Thus, even if the theory sets forth a necessary condition of the emergence of human ethical behavior—namely, that it in some way involves cooperation that preserves human genes—it does not set forth a sufficient condition of ethical ‘oughts.’”⁷ He elaborates:

The fuller understanding of this point requires some further consideration of just what kind of explanation is purportedly provided by evolutionary ethics. Some theorists construe it as a mechanical explanation, ultimately on par with physicochemical explanations. On this view, of course, the criticism I have just presented about the discontinuity with intentional ‘oughts’ would be reinforced. . . . This discontinuity, then, is one of the main grounds for raising the question: How ethical is evolutionary ethics? If to be ethical involves intending to act in certain ways that not only benefit other persons besides or in addition to oneself, but also are subject to knowledge, voluntary control, and reasoned choice on the part of the agent, then what evolutionary ethics presents as the content of what it calls ethical behavior is not, in fact, ethical. It is closer to a tropism than to a human action.

In other words, Gewirth’s objection rests upon the claim that biological accounts are insufficient to explain the content of agents’ values, for they do not come close to providing an adequate account of the cognitive-perceptual system that formulates, originates, and deliberates upon those values. While biological development might be necessary for such cognitive-perceptual and thus normative development, Gewirth’s important point is that it can be decisively shown that the lack of sufficient conditions as regards biological systems for normative systems beyond mere instinctual preferences undermines the direct connection evolutionary ethicists would like to make between normative value and the evolutionary legacy. As Gewirth argues in his own words:

When the evolutionary ethicist tries to account for the vast scope of the interests that must be subserved by moral conduct, he can do so only by stretching, beyond plausibility or empirical evidence, the reach of what he regards as biological causality. At a minimum, he here confuses necessary with sufficient conditions. The biological, evolutionary background is indeed the necessary condition of moral, and for that matter intellectual, aesthetic, and other cultural development. But it is also the necessary condition of immoral development; and it is not the sufficient condition of any of these modes of cultural development. And in trying to give an evolutionary explanation of moral judgments and moral conduct the evolutionary theorist lays himself open to the charge that he confuses the biological background without which morality cannot occur and the constitutive conditions which are the direct components of morality.⁸

⁷ Alan Gewirth, “How Ethical is Evolutionary Ethics?” *Evolutionary Ethics*. Ed. Matthew H. Nitecki and Doris V. Nitecki. (Albany: State University of New York Press, 1993): 245.

⁸ *Ibid.*, 253.

Gewirth thus lays bare the problems inherent in such purely inductive integrationist approaches.

There are yet other self-proclaimed scientific approaches to value theory which take an even more extreme stance against ethical rationalism. Although familiar to those immersed in the field of value theory or psychology, I will discuss them briefly. Prominent among such approaches are biological reductionism, eliminativism, and behaviorism. Although Rottschaeffer comprehensively addresses these competing accounts of moral agency and convincingly demonstrates their inadequacies, one additional point he does not fully address is the use of language within these different frameworks. Similar to some arguments against the reduction of psychology to neurophysiology, it would be utterly useless, let alone impossible, to reduce psychological accounts of moral agency either to epiphenomenal or neurophysiological accounts. The language we use to describe beliefs, feelings, desires, and other qualia involve the use of cognitive and folk psychological terminology. A reduction of these psychological phenomena to neurophysiological phenomena would, thus, involve an elimination of this vocabulary and, to a large extent, propositional attitudes. The lexicon of neurophysiology, having replaced the lexicon of qualia, would be inadequate to describe subjective experience and would effect a breakdown in the ordinary language of values and, thus, ethics altogether. The uselessness of such a reduction consists in the attempt to describe something readily observable in ordinary experience in terms of something that is not readily observable, as if subjective experience and the firing of neural networks were equivalent. Of course, there is an entire literature in Critical Theory, especially in the work of Habermas, and likewise in contemporary Analytic philosophy, devoted to the rebuttal of such claims. As such psychological theorists as Albert Bandura, and to some extent Lawrence Kohlberg, have demonstrated, agency and judgment are far more complex than can be accounted for by reference to notorious “attractors” and “repulsors.”⁹ The problems inherent in reductionism and eliminativism are important to point out before attempting a definition of rationality and agency, since, although highly relevant to certain aspects of meta-ethical theory, the vocabulary of neurophysiology often proves far less relevant than that of folk psychology and cognitive psychological theory in the analysis of the experience and execution of moral agency. The distinctly separate role of cognition, therefore, cannot be overlooked.

Individuals who take a biological approach to normative theory still might not be convinced. Indeed, while I have taken the approach of responding to progressively more and more sophisticated biological and evolutionary accounts of normative and ethical theory, one of my prime goals in this section, namely that of reconciling ethical rationalism and scientific integrationism, remains scandalously incomplete. I have not, as of yet, confronted what I take to be one of the most sophisticated arguments for an evolutionary account of ethical theory, that of Sharon Street.

⁹ For Bandura, specifically, see: Albert Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory*. (Upper Saddle River: Prentice Hall. 1985); *Self-Efficacy in Changing Societies*, (Cambridge: Cambridge University Press. 1995); *Self-Efficacy: The Exercise of Control*. New York: (W. H. Freeman and Company, 1997).

Although others have offered potentially devastating replies to her argument, these replies are mainly abductive and, while potentially powerful, are inherently not as powerful as deductive arguments.¹⁰ Without addressing her argument in a specifically deductive, aprioristic fashion, my efforts to demonstrate the aprioristic realism and intrinsicity of value might seem question-begging. In her notable paper, “A Darwinian Dilemma for Realist Theories of Value,” Street argues that value realism ultimately implies conclusions that are highly improbable. On the one hand, realism may implicitly assert a relation to the human evolutionary legacy and thereby force itself into the scientific realm as an explanatory hypothesis of the evolution of human values while simultaneously implying the highly unlikely coincidence that evolutionarily instilled values are those same ones that are independently sound or intrinsic or, on the other hand, it may deny such a relation and thereby imply the unlikely conclusion, due to the nature of human deliberation, that most of our evaluative tendencies are invalid. On both accounts, she argues, realist theories of value make highly improbable claims. Instead of going full force ahead into my rebuttal to this argument, I will simply begin by quoting Street’s final remarks in her paper and by making a perhaps very subtle point. She states: “Before life began, nothing was valuable. But then life arose and began to value—not because it was recognizing anything, but because creatures who valued (certain things in particular) tended to survive. In this broadest sense, valuing was (and still is) prior to value. That is why antirealism about value is right.”¹¹ While she seems to take a more modest view in that she allows for the phenomenon of “evaluative error,” I would agree, in some broad sense, with her implicit conclusions and with the implicit conclusions of all antirealist theories: there can be no definitive reason or reasons to believe that Street’s argument is sound. In other words, if Street’s argument is sound, then it follows that there is no “real” reason to believe it is sound. As a plethora of philosophers have aptly pointed out, inductive claims—indeed all claims—are not immune to normative regulation; reasons, including reasons to believe, are fundamentally normative.¹² My first point here is simply this: if, indeed, there is any reason to believe anything at all, Street’s argument included, and if such a reason holds whether or not there are sapient beings in the universe, then Street must abandon her line of argumentation, for her conclusions point precisely in the opposite direction. Street would likely respond that such an assessment is unfair, for there are indeed reasons to believe certain facts; it is simply

¹⁰ David Enoch, *Taking Morality Seriously: A Defense of Robust Realism*. (Oxford: Oxford University Press, 2011); Abraham Graber, “Medusa’s Gaze Reflected: A Darwinian Dilemma for Anti-Realist Theories of Value.” *Ethical Theory & Moral Practice*, 15 (2011): 589–601.

¹¹ Sharon Street, “A Darwinian Dilemma for Realist Theories of Value.” *Philosophical Studies*. 127 (2006): 156.

¹² Jean E. Hampton, *The Authority of Reason*. (Cambridge: Cambridge University Press, 1998); Alasdair MacIntyre, *After Virtue: A Study in Moral Theory*. (Notre Dame: University of Notre Dame Press, 1984); *Three Rival Versions of Moral Inquiry: Encyclopedia, Genealogy, and Tradition*. (Notre Dame: University of Notre Dame Press); Linda Trinkaus Zagzebski, *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. (Cambridge: Cambridge University Press, 1996).

that such reasons only came about, and could only come about, when there evolved valuers for whom they became reasons. In other words, there can be no reason to believe anything if there are no believers. Street, however, misses a crucial point: the truth of a proposition does not depend on whether or not there is an agent present to evaluate it. This, however, is what she implicitly rejects. If the phrase “reason to believe” seems problematic, we can reasonably transform it, for to assert that there is a reason to believe a proposition is to imply either that the proposition is true or that it is likely true. For instance, if reasons to believe or hold true certain facts arose only when perceivers of truth arose, then it follows that the truth of such a statement as “Light exhibits particle-wave duality” could not be said to hold when there are no agents. If this does not seem convincing, then consider that the following statement, call it proposition 1, must be true if there are indeed any reasons to believe anything at all: “There is no reason to believe something that there is no reason to believe.” If Street is right, however, the following statement must concurrently be true: “There existed a time t at which there was no reason to believe anything,” from which follows the statement “At time t , there was no reason to believe proposition 1.” From this it follows that “It is conceivable that, at time t , there was no reason to believe that the proposition ‘Proposition 1 is not-true’ was contradictory” and the lemma “The conceivability of there being no reason to believe that ‘Proposition 1 is not-true’ is a contradictory proposition at time t entails a claim on my part that I have a reason to believe this to be so conceivable.” Finally, from this it follows “I believe I have a reason to believe that it is conceivable that there is no reason to believe the proposition ‘There is no reason to believe something that there is no reason to believe’” or *a fortiori* “I believe it is conceivable that it is not the case that I do not have a reason to believe something that there is no reason to believe.”¹³ Street might reply that the argument is circular because it presupposes

¹³The *reductio* laid out more formally:

1. “There is no reason to believe something that there is no reason to believe.” (tautology)
2. “There existed a time t at which there was no reason to believe anything” (assumption)
3. “At time t , there was no reason to believe proposition 1.” (analytic consequence: 1, 2)
4. “It is conceivable that, at time t , there was no reason to believe that the proposition ‘Proposition 1 is not-true’ was contradictory” (analytic consequence: 1–3)

Lemma: “The conceivability of there being no reason to believe that ‘Proposition 1 is not-true’ is a contradictory proposition at time t entails a claim on my part that I have a reason to believe this to be so conceivable.”

5. “I believe I have a reason to believe that it is conceivable that there is no reason to believe the proposition ‘There is no reason to believe something that there is no reason to believe,’ and consequently, that I have a reason to believe that it is conceivable that I have a reason to believe a contradictory proposition is not a contradictory proposition.” (contradiction introduction, 2-4-4 lemma)

Restatement: “I believe it is conceivable that it is not the case that I do not have a reason to believe something that there is no reason to believe, and consequently that it is conceivable that I have a reason to believe that at least one contradictory proposition is not a contradictory proposition.”

Conclusion: Premise 2 is false. (analytic consequence: 1–5)

conceivability, that there must be a conceiver to conceive of the truth of “Proposition 1 is not-true.” This, however, is a poor objection, for what I am merely asking the reader to do is to attempt to conceive of a possible agentless universe in which “Proposition 1 is not-true” could be true. If it could not be true, in other words if there is never a reason to believe a fact that there is no reason to believe, then it simply *is* not true, irrespective of temporality. If it *is* not true, however, then the implications of Street’s and all anti-realist theories, namely that there are no reasons independent of agents, must be unsound. On this account, Street’s arguments concerning the dependence of normativity upon biology must be false. If Street’s argument is sound, by contrast, does it not follow that all of the good reasons to believe the theory of evolution too are merely grounded in the evolutionary process itself? This would seem to be a paradox of epic proportions. Street could bite the bullet, however, and state that truth is simply a socially or biologically constructed entity that depends on the context of human reasoning itself. Of course, she would then put herself in the position of denying the apriority of mathematics and logic, of for instance denying that planets *really* have circumferences in the absence of agents. She would then place herself in a potential dilemma between accepting *real* and *true* biological facts and accepting that truth is biologically constructed. Again arises the paradox. My statements here are by no means a definitive argument against Street’s claims; they are merely potential issues which I believe philosophers must consider before accepting Street’s claims or any claims asserting a fundamental justificatory role for a *posteriori* judgment.

My more substantive argument begins here. Let us begin with the first horn of this supposed dilemma. Street claims that value realists who assert a relation between the truth of value propositions and the basic set of evolutionarily developed values must accept a highly improbable coincidence, namely that humans incidentally evolved to know these truths. There is then the second horn of the dilemma: if value realists who deny this relation bite the bullet and claim that humans did not actually evolve to know these truths, then value realists must accept that all or most of our basic judgments are off track. As she states:

Of course it’s *possible* that as a matter of sheer chance, some large portion of our evaluative judgements ended up true, due to a happy coincidence between the realist’s independent evaluative truths and the evaluative directions in which natural selection tended to push us, but this would require a fluke of luck that’s not only extremely unlikely, in view of a huge universe of logically possible evaluative judgments and truths, but also astoundingly convenient to the realist. Barring such a coincidence, the only conclusion remaining is that many or most of our evaluative judgments are off track. This is the far-fetched skeptical result that awaits any realist who takes the route of claiming that there is no relation between evolutionary influences on our evaluative judgments and independent evaluative truths.¹⁴

Value realists, however, cannot thus claim that our judgments are off track because all of our judgments and deliberations about value presuppose the truth or at least the validity of some already inherited predisposition to value in some way. Our

¹⁴ *Ibid.*, 122.

cognitive and deliberative apparatuses are therefore already “contaminated” with basic instinctual values derived from our evolutionary legacy whose presupposed validity spills over even into our highly deliberate evaluations, including into our efforts to supersede or prove inaccurate our basic evaluative tendencies. In response to a proposed objection on the grounds that the argument ignores the reflective and deliberative capacities of human beings in formulating values and value systems, she argues:

The objection gains its plausibility by suggesting that rational reflection provides some means of standing apart from our evaluative judgments, sorting through them, and gradually separating out the true ones from the false—as if with the aid of some uncontaminated tool. But this picture cannot be right. For what rational reflection about evaluative matters involves, inescapably, is assessing some evaluative judgements in terms of others. Rational reflection must always proceed from some evaluative standpoint; it must work from some evaluative premises; it must treat some evaluative judgements as fixed, if only for the time being, as the assessment of other evaluative judgements is undertaken. In rational reflection, one does not stand completely apart from one’s starting fund of evaluative judgements: rather, one *uses* them, reasons in terms of them, holds some of them up for examination in light of others. . . Thus, if the fund of evaluative judgements with which human reflection began was thoroughly contaminated with illegitimate influence—and the objector has offered no reason to doubt *this* part of the argument—then the tools of rational reflection were equally contaminated, for the latter are always just a subset of the former.¹⁵

To clarify, the basic set of evolutionarily instilled values to which Street refers are sub-rational, what biologists and biological anthropologists who accept some version of the universal-values hypothesis identify in other animals as well. As she argues: “It is plausible to suppose that over the course of much of our evolutionary history what I have been calling ‘more basic evaluative tendencies’ were genetically heritable traits, where a *basic evaluative tendency* may be understood very roughly as an unreflective, non-linguistic, motivational tendency to experience one thing as ‘calling for’ or ‘counting in favor of’ something else. We may think of these as ‘proto’ forms of evaluative judgement.”¹⁶ We will revisit just these types of judgments periodically throughout this book with regards to social cognitive theory, the phenomenon of freedom, and the ideas of preference and desire. These phenomena are, I believe, important to distinguish from deliberative value judgment. Why this distinction is so important will become apparent as the arguments in this book proceed. For now and until such proof is given, I ask the reader to assume such a distinction here, as even Street herself has.

What shall we say to this supposed dilemma? In the first place, we may consider the peculiar idea that to assert a relation between the truth of certain fundamental value propositions and the evaluative tendencies instilled in humans by evolution entails the acceptance of a highly improbable claim about the coincidence of these evolutionarily instilled evaluative instincts and their validity. What, exactly, is improbable here? Street asks us to consider the wide variety of possible values

¹⁵ Ibid., 124.

¹⁶ Ibid., 119.

that could have arisen in humans but did not due to natural selection. As she states: “There are so many possible judgments about reasons we could make—so why these? Why, for instance, do we view the death of our offspring as a horror, rather than as something to be sought after? Why do we think that altruism with no hope of personal reward is the highest form of virtue, rather than something to be loathed and eliminated?” I believe we must ask a perfectly prudent question: what is actually included in the set of *possible* values humans could have acquired? Is it truly so large as Street makes it out to be? I think not. It is perfectly reasonable and consistent with evolutionary theory to state that these basic evaluative tendencies developed as they did because they made human survival and reproduction more likely. The important point to make here, however, is that there was no alternative; either organisms developed in this way and are now organisms for whom evaluation is possible in the first place, or no such organisms developed at all. These evaluative tendencies were, in essence, necessary conditions for the evolution of evaluative beings in the first place. Seen from this vantage point, it is not only *not* improbable that beings who valued survival would come to instinctually value survival but positively *necessary* that such beings value survival. Otherwise, we would not be speaking of beings at all. In effect, to have such evaluative tendencies might simply be what it means to be an evaluative *being* in the first place. In fact, as Philippa Foot has shown and as I will support and build upon in later arguments, such values as life and survival are fundamental parts of a grammar of value in general; in essence, one cannot speak of value without invoking *existence* and *life*.¹⁷ If the invocation of value just does imply the invocation of existence and life, then there is no problem of improbability or unlikely coincidence here. Thus if altruism, the care of one’s offspring, and the like, are indeed valuable, then there is no conflict here with evolutionary theory. Instead, we may simply say that to have these evaluative tendencies just is a manifestation and perpetuation of the phenomena of value and evaluation themselves. Of course, this gets us squarely into one of Street’s main objections to what she calls “tracking accounts” of value, namely that such accounts constitute scientific hypotheses and thereby compete with existing scientific theories of evolution. The explanation of value that I have set forth, however, is explicitly *not* a scientific one and does not pretend to be one. Instead of claiming that humans came to have the basic evaluative tendencies they did because, as Street states, “they are true,” I unequivocally admit that evolutionary theory explains why we have such evaluative tendencies *but* that evolutionary theory does not and in principle cannot explain why such evaluative tendencies have value. Of course, we may yet question the precise listing of evaluative tendencies that Street has laid out. In the first instance, it seems as though she has assumed that all realist theories of value come to roughly the same conclusions about what is valuable. That is, however, not the case. We do not need to explore the differences between the various theories of value which exist, but needless to say, there are many existing theories which are fundamentally at odds with many or most

¹⁷ Philippa Foot, *Natural Goodness*. (Oxford: Oxford University Press, 2001).

people's ethical instincts and intuitions. At this point, Street need only point once again to the "contaminated" nature of the deliberative route with respect to evaluation in order to show, she thinks, that all derived evaluation is subsumed under the basic evaluative tendencies humans inherently possess. This objection, however, only strengthens my own argument, for I can once again reply to Street, "But how is evaluation and value itself conceivable without the presupposition of certain specific initial values?" While I referred previously to the conceptual-linguistic necessity of life and survival to the concept of value, I here propose, much inspired by Habermas, that discourse, imagination, and creativity, and implicitly the conditions for their possibility and realization, embody the concept of value as well in that, once again, the concept of value and evaluation is impossible without them and, most importantly, without pragmatically asserting their value; this will be proved later. Thus, it does not matter that the deliberative route is "contaminated" by certain biological evaluative tendencies, for such "contamination" is what makes the concept of value intelligible and meaningful in the first place. Recall that this argument is not so far off from what Richards himself has argued, although as I have claimed, the foundations of his model are inherently conflictual. The outcome of my rebuttal follows thus: while, as Gewirth has already argued, natural selection constitutes the necessary conditions for evaluative judgment, value is not reducible and in principle cannot be reduced to biology, for it is not dependent on biology or evolution as chemistry, for instance, is upon physics. Those necessary conditions which have happened to give rise to evaluative beings do not determine arbitrarily what it is that, as evaluative beings, we must value. Rather, it is simply the case that natural selection gave rise to evaluative beings whose definition includes, by necessity, those things which evaluative beings value qua evaluative beings and those things which are valuable qua the concept of value. Thus, it is not a fact of evolution that evaluative beings possess certain tendencies, for to be valuers and to possess any values, there are certain tendencies an intelligent being *must* have, whether it be a human, an AI, an alien, or an angel, and who is to say how any of these non-human entities would come about? We must therefore reject the idea that value and its conceptual links to certain associated behavioral tendencies depend upon biological contingencies instead of upon aprioristic foundations. We may reject these tendencies if we so choose, and we may even decide that some of them are off the mark, but we will do so at the risk of pragmatic inconsistency. As for the many other derived values which people may value, such as the color purple, the texture of silk, or the smell of gasoline, we may refer, once again to our most basic evaluative tendencies, working our way through our higher cognitive and deliberative processes, and we may either accept or reject them after such reflection, but the most basic evaluative tendencies, those which aid us in surviving and evaluating alike, are literally not capable of being evaluated away.

Thus, we now find ourselves in the position of clarifying some of the misrepresentations which have, perhaps, caused such integrationists as Rottschaeffer and Hauser to shy away from ethical rationalism. There have, most certainly, been major problems with rationalist ethical theories in the past, and many of these approaches have been stigmatized by such theorists as Hare as

“Cartesian” in nature.¹⁸ This insult, no doubt, was aimed primarily at Kant. While Kant’s ethical theory has been seminal to the growth of various ethical and meta-ethical insights, its various flaws have seemed to cause a peculiar revulsion on the part of particular theorists to deontological ethical rationalism in general. The problems most prominent in Kant’s ethical system include the supposed non-hypothetical justification of the categorical imperative, the justification of his conception of universalizability, and his distinction between autonomy and heteronomy. It is this last issue which has, perhaps, been most problematic for the relationship between ethical rationalism and ethical naturalism, for not only does Kant’s conception of autonomy hinge on the non-natural or preternatural, but the autonomy-heteronomy distinction, as set forth in Kantian and some neo-Kantian theories, has proved to be mired in contradictions and altogether inaccurate conceptions of agency. In recent years and in the modern period more generally, a whole host of philosophers have come forth to address these problems, and often with promising results.¹⁹ We must realize, therefore, that while Kant’s work has undeniably influenced the shape and direction of all subsequent work in deontological rationalism, his theory has undergone so many reconstructions since he first put pen to paper that many of the models that have emerged from such reconstructions share, we might say, only a family resemblance to Kant’s original work. In the first place, rationalist ethical theories often have a strong naturalistic component, as is the case with Gewirth. His Principle of Generic Consistency centers on what he

¹⁸ R. M. Hare, *The Language of Morals*. (Oxford: Oxford University Press, 1952): 39.

¹⁹ H. Allison, *Kant’s Theory of Freedom*. (Cambridge: Cambridge University Press, 1990); Karl-Otto Apel, *La Réponse de L’Éthique de la Discussion au Défi Moral de la Situation Humaine Comme Telle et Spécialement Aujourd’hui*. (Louvain: Éditions de l’Institut Supérieur de Philosophie Louvain-la-Neuve, 2001); Marcia Baron, “Virtue Ethics in Relation to Kantian Ethics: An Opinionated Overview and Commentary.” *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics*. Ed. Lawrence Jost and Julian Wuerth. (Cambridge: Cambridge University Press, 2011): 8–38; Brand Blanshard, *Reason and Goodness*. (Atlantic Highlands: Humanities Press, 1961); Rainer Forst, *The Right to Justification: Elements of a Constructivist Theory of Justice*. (New York: Columbia University Press, 2012); P. Guyer, *Kant and the Experience of Freedom*. (Cambridge: Cambridge University Press, 1993); *Kant on Freedom, Law, and Happiness*. (Cambridge: Cambridge University Press, 2000). Jürgen Habermas, *Moral Consciousness and Communicative Action*. (Cambridge: The MIT Press, 1990); T. Hill, *Dignity and Practical Reason in Kant’s Moral Theory*. (Ithaca: Cornell University Press, 1992); P. Kleingeld, “What do the Virtuous Hope for? Re-reading Kant’s Doctrine of the Highest Good” in H. Robinson (ed.), *Proceedings of the Eighth International Kant Congress*, Milwaukee: Marquette University Press 1: 91–112 (1995); Christine Korsgaard, *The Sources of Normativity*. (Cambridge: Cambridge University Press, 1996); Onora O’Neill, *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. (Cambridge: Cambridge University Press, 1989); *Towards Justice and Virtue: A Constructive Account of Practical Reasoning*. (Cambridge: Cambridge University Press, 1996). Derek Parfit, *On What Matters: Volume I*. (Oxford: Oxford University Press, 2011); *On What Matters: Volume II*. (Oxford: Oxford University Press, 2011); John Rawls, *A Theory of Justice*. (Cambridge: Harvard University Press, 1971); T. M. Scanlon, *What We Owe to Each Other*. (Cambridge: Harvard University Press, 2000); Marcus George Singer, *Generalization in Ethics*. (New York: Alfred A. Knopf, Inc., 1961); R. Sullivan, *Immanuel Kant’s Moral Theory*. (Cambridge: Cambridge University Press, 1989).

calls the “normative structure” of agency.²⁰ As such, it addresses ethics in a naturalistic way, instead of in a sort of preternaturalistic way, and thereby avoids the problems inherent in the autonomy-heteronomy distinction. From the work of Gewirth, we can see how a rationalistic, deontological theory can be intricately connected to naturalism and, perhaps more importantly, how it might have great relevance to neurophysiological and psychological accounts of agency when approached from an integrationist, evidence-based standpoint. Ethical rationalism, then, does not imply non-naturalism, preternaturalism, or a complete lack of consultation with empirical reality, nor does it imply adherence to the Kantian autonomy-heteronomy distinction. A failure to acknowledge this, and to class all deontological or rationalist theories as Kantian or, otherwise, pseudo-Kantian, has led a host of philosophers to preclude a whole set of rationalist and deontological theories, namely dialectical theories, from consideration. While most scientific integrationist arguments are founded upon inductive syllogisms comprised of empirical premises, and while many rationalist arguments are founded upon deductive syllogisms comprised of excessively formalistic non-empirical premises, Gewirth’s theory, as well as a Gewirthian-integrationist synthesis, yields a deductive syllogism founded upon empirical premises. Thus, Gewirth’s dialectical method, when additionally supplemented by biological and psychological evidence, demonstrates that a rationalistic, deontological theory can be both *a priori* and grounded firmly upon empirical facts. Theories which have this structure, as we shall see, do not run into the problem inherent in such accounts as those of Richards or his colleagues either.²¹

Hare’s theory of preference-utilitarianism, while also a rationalist theory, is a non-cognitive, anti-naturalist, and anti-descriptivist one. While it would be wrong to group Hare together with intuitionists, emotivists, and preter-naturalists, and while it would certainly be erroneous to discount his theory completely, his position, as it stands, has only limited application. As I will demonstrate in further sections, Hare’s theory yields guidelines of permissible action which pertain only to commonly held preferences, instead of, additionally, to momentary, uncommon, or eccentric preferences, as Hare would like to claim. The primary importance of Hare’s theory, instead, is twofold. First, his theory offers a bridge to scientific integrationists by demonstrating to them that ethical theories based on the biological and psychological phenomena of altruism, empathy, preference, and desire provide at least a valid basis for ethical reasoning and an alternative to dead-end theorizing from *a posteriori* premises alone. If the part of Hare’s theory that I

²⁰ Gewirth, *Reason and Morality*, 26.

²¹ Richard D. Alexander, “Biological Considerations in the Analysis of Morality.” *Evolutionary Ethics*. Ed. Matthew H. Nitecki and Doris V. Nitecki. (Albany: State University of New York Press, 1993): 163–196; Andrzej Elzanowski, “The Moral Career of Vertebrate Values.” *Evolutionary Ethics*. Ed. Matthew H. Nitecki and Doris V. Nitecki. (Albany: State University of New York Press, 1993): 259–276; Lawrence Slobodkin, “The Complex Questions Relating Evolution to Ethics.” *Evolutionary Ethics*. Ed. Matthew H. Nitecki and Doris V. Nitecki. (Albany: State University of New York Press, 1993): 337–347.

highlight is sound, then it could provide a useful bridge to Universal Moral Grammar theorists, although Hare might cringe at this. Second, his theory successfully demonstrates that at least some basis for resolving ethical disputes is possible, which will be important when I examine the possibility of a method of justification. Although Hare's might not be the fundamental or the most sound basis, it at least invalidates many relativist, subjectivist, and nihilist claims. As we shall see, it is Gewirth and Habermas's theories which provide the fundamental basis for moral judgment (although not value judgment) and which, through their implicit alliance with naturalism, invalidate most relativist, subjectivist, and nihilist claims.

2.1.2 Reasserting the Compatibility of Virtue Ethics and Deontological Rationalism

Virtue ethics has also been framed as irreconcilable with and, indeed, the enemy of deontological ethical rationalism. The foundations of this view can be found most prominently in the work of G.E.M. Anscombe.²² Macintyre, probably the most vocal proponent of this view, has argued that deontological theory, as well as utilitarian theory, is a misconceived and misguided project doomed to failure because of its Enlightenment roots. This ultimate failure, he asserts, is the result of Enlightenment thinkers' off-hand dismissal of Aristotelian theory. This rejection, he claims, has had its ultimate culmination in the misguided theories of emotivism and intuitionism. The claim that Enlightenment thinkers dismissed without argument the traditions of Aristotelianism and Scholasticism is, for the most part, uncontroversial. It is well known that Enlightenment thinkers, in general, abandoned formal and final causality, as well as most classical conceptions of virtue. As a result, Enlightenment ethics has, admittedly, traversed a misguided path, but it has not all been misguided. Enlightenment thinkers, in fact, filled in conceptual gaps where various classical thinkers seem to have overlooked them. Deontology and virtue ethics, in fact, share a common root, one that will be explored in depth in the course of my argument for an assertorically necessary, post-dialectical theory of value. Macintyre's attempt to divorce himself entirely from the legacy of the Enlightenment, however, has led him to reject any form of deontology, and many have followed his lead. While the concepts of formal and final causality, as well as the essence-accident distinction, are worth reincorporating into ethics, Macintyre's total rejection of ethical rationalism, defended mostly by some valid criticisms of Kant (whose theory I previously contended has been, in any case, variously reconstructed) and some very shallow criticisms of Gewirth, is unwarranted. Furthermore, his characterization of the Enlightenment as largely defined by a supposedly fallacious rational foundationalism is, I believe, inaccurate. It is, perhaps, his

²² G. E. M. Anscombe, "Modern Moral Philosophy." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. (Oxford: Oxford University Press, 26–45. 1997).

near-exclusive focus on particular aspects of the virtue theories of Plato and Aristotle, as well as his emphasis on cultural and traditional narrative, which leads him to neglect such foundationalism in the ancient world. Classical scholar, John Lenz, suggests that such mistakes are pervasive. He explains:

A . . . threat today is the axiom that culture trumps nature. . . Epicurus, like most ancient Greek philosophers, believed that wisdom was outside of and above politics, in understanding the universe and ourselves, together. . . Rational knowledge, *accessible to all human beings through our minds*, provides the best means of living well. . . This is not the popular Aristotelian practical wisdom.²³ (italics mine)

The rejection of anti-foundationalism, despite such thinkers as Richard Rorty's urging to the contrary, can only lead to progress in knowledge and not, as Rorty suggests, to a merely communal truth.²⁴ Appeals to the supposed primacy of culture, tradition, or history, even if they do turn out to be in some sense primary, do not change this.

Notably, virtue ethics takes a completely different approach than its Enlightenment cousins: while such models as deontology, utilitarianism, emotivism, and intuitionism ask the question "What should I do?," virtue theorists ask the question "What kind of person should I be?" and thus implicitly invoke the kinds of questions I posed at the outset. While many deontological and rationalist theories might be incompatible with the latter question due, as MacIntyre rightly claims, to these theories' failure to acknowledge a teleology, it does not follow that they all are. While the connection between such approaches as scientific integrationism and Aristotelianism might be clearer, since they both acknowledge the relevance of functionality statements to value judgments, the connection between Aristotelianism and deontological ethical rationalism might, at first, be more difficult to spot. In order to see the connection more clearly, we must pick out some of the formal components of virtue ethics. Although the primary focus in such a model is on what kind of person one should be, virtue ethics does, indeed, implicitly pose the question "What should I do?" for the kind of person one is clearly has direct bearing on what kinds of things one should do, whether or not the idea of *being* and *doing* are conceived in the same manner. To understand this, let us analyze a virtue theorist's response to the question "What should I do?" The answer, of course, would not be anything resembling "Follow rule A," since most classical conceptions of virtue ethics reject any such notion of rules. Rather, the response would be "Do what a virtuous person would do." Although this seems like a truism or a glaring generality, this response is actually quite telling. What this answer reveals, and what proponents of virtue ethics acknowledge, is that the learning, practice, and ultimate habituation of virtuous behavior is integral to the making of a virtuous person. In other words, an essential part of being virtuous is the development of a natural disposition to be virtuous, and thus virtue cannot but begin with a profound

²³ John R. Lenz, "How Epicurean Science Saves Humanity in Lucretius," *Lucretius: His Continuing Influence and Contemporary Relevance*. Ed. Madigan & Suits. (Rochester: RIT Press, 2011): 92.

²⁴ Richard Rorty, *Philosophy and the Mirror of Nature*, (Princeton: Princeton University Press. 1979).

and honest self-appraisal. But how do we form such a disposition, and how might we undertake such self-appraisal? Deontological rationalism provides us with initial hints.

Deontological ethical rationalists assert that principles of action must be established before any pattern of behavior can be endorsed. Although Macintyre attempts to argue himself away from this line of thinking by asserting that patterns of behavior are inextricable from the culture in which they operate, this, by no means, obliterates the concept of moral principles. In addition, although he argues against the concept of moral rules, he does not, however, argue against the necessity of principles of behavior *per se*, as he and various virtue theorists seem to wish. Rather, he argues only against rules, not principles, of behavior, and this distinction is vital. A central virtue, he asserts along with Aristotle, is *phronesis*, or the ability to judge well in particular circumstances. Since moral rules cannot bend to the peculiarities of particular circumstances, he claims, it is instead necessary for a person to exercise *phronesis*. In order to demonstrate this, he gives the example of a land dispute for which there is an ongoing lawsuit. As he states: “Rule-specified concepts of justice. . . can give us no help at all. . .” since “the problem. . . concerns a period of time in which we do not as yet know *either* who has a just title by acquisition and transfer, for precisely that is to be decided by the current legal case.” He uses this example to argue that “rough and ready reasoning” guided by *phronesis*, and not moral rules, is the only adequate way to address such situations. As he states:

For each virtue therefore there are two corresponding vices. And what it is to fall into a vice cannot be adequately specified independently of circumstances: the very same action which would in one situation be liberality could in another be prodigality and in a third meanness. Hence judgment has an indispensable role in the life of a virtuous man which it does not and could not have in. . . the life of the merely rule-abiding man.²⁵

While principles in most models of ethical rationalism are what one might refer to as unchanging, absolute, or even “rigid,” we should expect the rules derived from them to be situationally dependent. For the sake of argument, take the principle “Preserve others’ health.” This principle, when applied to different situations, will lead to an assortment of different rules. In one situation, you might preserve a person’s health by giving him or her antibiotics, while in another situation, you might do it by feeding him or her vegetables. Naturally, rules such as “Give antibiotics to a person with bacterial meningitis, as long as some conditions X, Y, and Z are also met” and “Feed your children vegetables instead of candy, as long as some conditions X, Y, and Z are also met” are derived, yet these rules both satisfy the supposedly rigid principle to preserve others’ health. We might acknowledge how relevantly similar, at least in form, this principle is to such implicit principles of behavior as “Cultivate virtue within yourself” or “Do good and avoid evil.” The point is that, in order to exercise *phronesis* at all, we must base our judgment on

²⁵ Alasdair Macintyre, *After Virtue: A Study in Moral Theory*, (Notre Dame: University of Notre Dame Press. 1984): 154.

some initial behavioral principle, or at least some general goal, whether it be *a priori* or *a posteriori*. Certainly, if we are to abide unconditionally by the principle “Preserve others’ health,” we would not be justified in regularly feeding our children candy or in depriving a person with bacterial meningitis from the appropriate antibiotics. To do this, in fact, would be an exercise of poor judgment. In other words, we would be exercising a vice, although Macintyre and others rightly acknowledge that virtues and vices are not, strictly speaking, rule-based. If we adhere to the same pattern with “Cultivate virtue within yourself,” what this demonstrates, at the very least, is that principles of behavior are compatible with the exercise and attainment of virtue and, possibly, that virtue manifests itself through principle. This is a point Rosalind Hursthouse and Michael Slote have also suggested.²⁶

The development of a natural disposition to be virtuous through habituation can only occur if there is some end to which the virtues function, namely the *summum bonum*, and on this point, Aristotle and other classical philosophers would generally agree. The seeming rift in contemporary debates on virtue ethics, however, is partly the result of the complete alienation of judgment and rule from each other. When a person exercises *phronesis*, he or she is making judgments using practical reason. In order to make such judgments, however, one necessarily considers the relevant criteria of a situation, or, as Karen Stohr puts it, one exercises a “capacity for ascertaining saliences peculiar to the given situation.”²⁷ One must formulate, then, even if momentarily, some mode of action guided by some method of practical reason. This, even if in some rough sense, constitutes a rule-guided decision, unless Macintyre is speaking of rules, themselves, as rigid and inflexible non-teleological imperatives; but if this is how we are to interpret the concept of a rule, then, certainly, we must throw out the concept of rules. If we are to dispose of the conception of rules which I have suggested, however, namely rules as imperatives aimed toward a *telos*, it seems that the only mode of decision-making left for the virtue theorist is an emotivist or intuitionist approach, a specter from which Macintyre has insistently and rightfully estranged himself. These modes of decision-making, of course, are not the only ones left open to the virtue theorist; all decision-making, as we shall see, is teeming with, although not founded upon, rules.

If we are to bridge this unnecessary gap between judgment and rule, and if we are, additionally, to acknowledge that Macintyre has not truly attacked moral principles as such, we might begin to see that virtues, principles, and rules are not mutually exclusive. We might see, rather, that they each play a distinct role in voluntary action and that, conceptually, they are inextricably interwoven. It is plainly conceivable that the *summum bonum*, if there is one, might provide the

²⁶ Rosalind Hursthouse, “Virtue Theory and Abortion,” *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. (Oxford: Oxford University Press, 217–239. 1997); Michael Slote, “Agent-Based Virtue Ethics,” *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. (Oxford: Oxford University Press, 239–236. 1997).

²⁷ Karen Stohr, “Contemporary Virtue Ethics,” *Philosophy Compass*. 1, no. 1. (2006): 25.

foundation for a supreme moral principle or principle of conduct, for, although Gewirth and Habermas do not explicitly acknowledge it, we can see in both their and many aretaic models a teleology based upon the agent's constitution. We might, then, see that principles of behavior are regulated by the underlying function of the virtues and of aretaic concepts in general and, in turn, that such principles, themselves, are actually a manifestation of underlying aretaic concepts. It seems that we must ask an important question: "When we finely examine the nature of such concepts as *to ti en einai*, *ergon*, *telos*, *ataraxia*, and *eudaimonia*, do we actually discover a vital and well-defined link between aretaic and deontic concepts?" The answer, I believe, is yes. What we find is a system of deontological rationalism, but one that is wholly founded upon and answerable to *arete*, *telos*, and intrinsic goodness. This is a momentous departure from most deontological frameworks, whose proponents argue that aretaic concepts are illegitimate unless they are founded upon solid deontological grounds. On the contrary, both systems function together to yield a powerful account of ethical theory, but, despite the claims of deontological theorists, it is deontology, and not virtue ethics, that cannot stand on its own. This is especially true as regards the concept of *to ti en einai* or essence, implicit in the idea of *telos*, without which the inculcation of virtue via the necessary route of self-appraisal would be impossible. It might not be obvious at the outset why the idea of self-appraisal and subsequent self-analysis is a vital component of virtue, but we shall see in later chapters that it forms a vital part of a grammar of goodness and particularly of betterness; not only must *telos* be reincorporated into value theory, but so along with it must its correlate, *to ti en einai*. Going forward, it will become increasingly clear that, while deontological rationalism, scientific integrationism, and virtue ethics all elucidate specific pieces of the ethical puzzle, it is necessary that their mutual alienation from each other cease and their necessary integration, what Stohr endorses as "important trends," be recognized.²⁸ Such an integration, although considerable, is not simply a *via media*; it is not, in other words, a mere philosophical exercise in imaginative thinking. Rather, it should be clearly noted that this integration is the result of the recognition that deontological rationalism fails to yield for us a full and complete understanding of value theory (and even of ethical theory more specifically) and that virtue theory, while currently fraught with a certain degree of indeterminacy and vagueness, is indeed capable of filling these theoretical gaps via a reconstruction and re-appraisal of its essential features. Although this section has merely served to probe the two theoretical paradigms of deontological rationalism and virtue theory and to suggest loose ends, so-to-speak, in prominent modern views on their relationship, I will, as this book continues, press even harder in presenting formal challenges to these prominent modern views, demonstrating decisively that deontological rationalism, as well as other Enlightenment-born models, do not and cannot in principle have application to certain fundamental problems and questions of value where virtue theory, conceived and reconstructed on certain grounds, exhibits facility in doing

²⁸ *Ibid.*, 26.

so. As such, I do not negate deontological rationalism but, on the contrary, provide a foundational support for it as a powerful and clearly specified supplement to the aretaic concept that I will claim *is* value. In the meantime, and as our next step, let us proceed to further unravel some ornery loose ends.

2.2 The Failure of Axiological Anti-foundationalism

In this section, I problematize anti-foundationalist approaches to value theory and demonstrate the necessity of taking an overtly foundationalist approach. I argue that anti-foundationalist or non-foundationalist ethical and value theories simply lack any prescriptive or directional force. As such, they flounder in providing resolutions to, or even in providing reliable methodologies for addressing, moral problems and dilemmas of significant practical relevance.

2.2.1 *The Ideological Dangers of Anti-foundationalism*

Before we progress in our analysis, we must first address a controversy that lingers in value theory concerning the supposed dubiousness of rational foundationalism, or the derivation of theory from a fundamental, logically necessary principle or set of principles. While nearly all proponents of the anti-foundationalist position claim that foundationalism is inherently a flawed or failed undertaking, leads only to confusions, and can never yield any significant ethical truths, what I would like to argue is that this is precisely the state of anti-foundationalism. While anti-foundationalists typically abandon rational foundations for the very purpose of seeking and providing a clearer understanding of value theory, time and time again, all anti-foundationalist approaches have done is to further confuse value theory. Writings which follow this approach are characterized by a few easily identifiable features. The first feature is a set of arguments setting forth some significant intuitive, social, or biological phenomena and establishing them as supremely relevant and, thus, justificatory and prescriptive. These phenomena might include anything from the prevalence of empathy and altruism in moral decision-making, various game-theoretical, rational-decision-theoretical, or social-choice-theoretical models to moral intuitions, cultural constructions, or genealogy. The incontrovertible reasons for the relevance of such factors is almost always absent, and any epistemological explanation of how we would discover the incontrovertibility of these reasons is also notably lacking. Instead, we are often made to infer that these reasons are somehow self-evident, simply incontrovertibly apparent, or, best of all, “rational.” Another feature of these philosophers’ writings is somewhat predictable: because such an approach is to some extent inherently unmethodical and based fundamentally (whether they would like to acknowledge it or not) upon a slew of assumptions and prejudices, their main works on the topics of ethics, justice, or value in general almost invariably

contain material that lacks a certain degree of directional force or coherence. We are often left with a great many good ideas, but we are lost as to how we are to incorporate and systematize these ideas so that we may answer or approach answering such questions as “Is voluntary active euthanasia permissible?,” “Is labor-capital collective bargaining obligatory on the part of capital?,” or even “Is slavery wrong?” My point here is not to deny that science or social theory are important to ethics—they are. My point is merely that the simple fact that the amygdala is associated with emotionally influenced decision-making, that people act in a self-interested manner in situation X, or that my cultural heritage is highly relevant to those virtues I choose to cultivate cannot explain anything substantive concerning ethical theory *by themselves*. Instead, they need to be incorporated into a framework of reasoning which is grounded, yes, upon some basic, necessary, logical foundations. Even philosophers who are not typically thought of as strictly anti-foundationalist often fall prey to anti-foundationalist tendencies. Although some percentage of the population may share the moral intuitions of such philosophers as Parfit, Scanlon, Nagel, Rawls, Pogge, Dworkin, and the like, and although another percentage of the population may share the moral intuitions of Nozick, Gauthier, Hoppe, Narveson, Rothbard, and the like, or even of Kropotkin, Bakunin, Proudhon, Goldman, and the like, these thinkers should not build theories so as to constrain “rational” positions to those which fit only their own intuitions, which, more often than not, they mistakenly identify as universally shared. In commenting on Gauthier’s theoretical work, for instance, Parfit states: “Such conclusions, Gauthier concedes, conflict strongly with most people’s moral beliefs. But Gauthier rejects appeals to such intuitive beliefs, or to our ‘considered moral judgments’, which he claims that moral theories ought to ignore.”²⁹ Gauthier’s attitude, however, is reasonable, for unless moral intuitions can be shown to have some relevance to moral theory (which they may well have), we simply are not justified in appealing to them for the establishment of our case. Of course, it is conceivable that there is some principled and reasoned link between moral intuitions and moral justification, but this link, if it exists, has not been elucidated. Most notably, when Parfit wishes to attack or denounce some moral view, he often refers to the position’s intuitive implausibility or to its irrationality. Indeed, his entire reformulation of the categorical imperative and other Kantian precepts is shaped by what seems, to him of course, intuitively implausible. Although not truly anti-foundationalist, his account, as in anti-foundationalist accounts, seems merely to appeal to what seems self-evident or intuitively most plausible. I believe this to be a dangerous course in the establishment of value theory. There are notably deep divisions among philosophers, for instance, concerning the existence of positive rights. Obviously, then, there are those who do not share the intuition that such rights exist. To put it more bluntly, such philosophers as Rawls, Nagel, Scanlon, Parfit, Pogge, and Dworkin seem to base the plausibility or rationality of any ethical theory upon certain state-left-leaning standards and judge any formulation of ethical principles against such standards. In other words, such philosophers simply assume that

²⁹ Derek Parfit, *On What Matters: Volume I*. (Oxford: Oxford University Press, 2011): 346.

“moral intuitions” in general are those of a social democrat or a left liberal, as if no other intuitions about morality and political theory existed. They thus act as if it were a universally held intuition that states should develop welfare systems for the poor, including health care, education, and guaranteed employment, and that they should also develop complex systems of taxation. This, however, is an absurd assumption, and it exposes the partial truth of Nietzsche’s famous proclamation: “Mankind does not strive for happiness; only the Englishman does that.”³⁰ People of various different moral persuasions, it can be seen, have and have historically had wildly different moral intuitions. Indeed, there are fascists, state socialists, state communists, libertarian socialists, libertarian communists, state capitalists, libertarian capitalists, monarchists, plutocrats, aristocrats, racists, sexists, homophobes and transphobes, classists, and other types of elitists. Of course, their standards of theoretical plausibility and rationality would also presuppose the soundness of their own positions. Any one of these theories may be right, including Parfit and his colleagues’ own respective interpretations, but this has not been proved, at least not by any non-circular kind of reasoning. While Parfit should be commended for his thorough and exhaustive work in ethical theory and the great fruits it has produced, it seems that his theory, too, lacks fundamental justifiability. Of course, while we do not all share the same moral intuitions, we might all share the same fundamental values, but this view suffers from problems as well.

As we saw in previous pages, scientific integrationists such as Hauser and Rottschaeffer rely largely on Universal Moral Grammar Theory, which posits a basic, shared, universal code of ethics and sympathetic-empathetic reasoning. The simple fact that all people share a basic foundational code of ethics and certain basic values cannot, however, speak to which of these values, if any, we *ought* to value. Likewise, the simple fact that many of our moral judgments rely on sympathy or empathy cannot provide any substantive prescriptions concerning *how* we are to use these faculties, if at all.

We already saw, as well, that Macintyre’s views, while not nearly representative of all virtue theory, end up largely in the same place. From none of his works can we glean what it is we are most justified in doing, besides perhaps reasoning and social interaction in general. Because of his ambiguous use of the concept of virtue, as well as his failure to systematize an explanation of exactly how we are to derive virtuous behavior from our respective histories and cultural traditions, we are left wondering how we will ever answer the important, pressing questions which face us in the world.

Thus, what we observe is the principal danger that anti-foundationalism yields: the creeping specter of ideology. By assuming some observation to be intuitive or in some way obvious, without providing explicit proof, we always already presuppose some foundation, whether we like it or not. This foundation, however, is an insidious one, for it goes un-analyzed and unexamined, positing itself as somehow

³⁰ Friedrich Nietzsche, *Twilight of the Idols: or How to Philosophize with a Hammer*. (Oxford: Oxford University Press, 2009).

not in need of justification. Although it is perhaps not fair to state that the aforementioned theorists are steeped, so-to-speak, in ideology, the presumption of intuitive self-evidence or of the otherwise obvious relevance of certain factors or phenomena to ethical judgment, has the capacity to lead in this direction.

2.2.2 *The Problem of Indeterminacy: Amartya Sen's Theory of Justice*

The most recent rejection of foundationalism is also perhaps the most well-argued. In *The Idea of Justice*, Amartya Sen utilizes social-choice theory and various other constructs from set theory in order to construct a theory of justice. Despite his thorough approach, however, he, like all other anti-foundationalists, bases his premises on what I believe are merely prejudices and assumptions concerning what is just and unjust. Transcendental institutionalism, as he terms a particular approach to justice-theoretical foundationalism, is neither necessary nor sufficient for a theory of justice, for we can supposedly make judgments concerning what is “manifestly unjust” without making a judgment concerning what is the most just possible state of affairs. As we can compare the tallness of two mountains without making reference to a tallest mountain, we can make judgments of the sort “Freedom is better than slavery” without knowing what the most just possible arrangement is within a free society.³¹ Certainly, this works in the case of comparing wide things, tall things, short things, and the like, but he seems to overlook an entire legacy of meta-ethical analysis which seems to demonstrate that the comparison of good things, bad things, and the concepts of goodness and evil themselves, do not conform to standard grammatical conventions concerning predicativity and attributivity, at least not in the *exact* way that these other sorts of adjectives do. As I will clarify in later chapters, goodness is what I shall call *reflexively intrinsic* and is constituted by certain interrelated properties which give to the property of goodness a slightly different function than other intrinsic properties. One of the relevant concepts concerning goodness comparison, or what I shall call betterness for shorthand, that of *telos*, is vital to understanding value comparison. The importance of the re-emphasis of *telos* is perhaps one of the greatest contributions to ethics that Macintyre has given us, and it bears emphasizing here that Sen, as well as many others, regularly overlook the importance of the concept of *telos*, as well as *to ti en einai* or essence. As a result of this, Sen makes the false assumption that certain values are incomparable; he does not, of course, consider that two values can *always* be compared to each other if we regard such comparison as, for instance, a comparison of each value's tendency to fulfill a given *telos*. Good things are, indeed, impossible to compare without invoking *telos*, and goods can only be compared by observing their respective statuses relative to a

³¹ Amartya Sen, *The Idea of Justice*, (Cambridge: Harvard University Press. 2009).

given *telos*; this position will be defended at length elsewhere. Social-choice theory can tell us a lot about the interaction of various values, but it tells us nothing about how valuable one state of affairs is in comparison to another. While Sen tells us that certain states of affairs, such as famine, are manifestly unjust, he is merely implicitly asking us to assume that this is so; he provides no concrete proof to demonstrate the truth of this claim. This might sound like a callous assertion at the outset, tantamount to denying the evil of famine, war, and other comparable states of affairs, but this is not so. I fully acknowledge the evil of these states of affairs, but my point is merely this: because normative assumptions are so pernicious and have been historically so pervasive and deeply ingrained in our minds (with what I believe most will agree have been disastrous consequences), we must sweep away all normative assumptions and begin from scratch. Let us not assume anything, even if it “feels” correct or “is self-evident.” It may “feel” correct that famine and war are evil or unjust, but it is by no means “self-evident” that they are. If they are, there must be incontrovertible grounds for why they are, and if they are, we must be convinced by thorough and rigorous proof. It is, it seems, one of Sen’s prime motivations for constructing an anti-foundationalist theory that we will supposedly forget all about the practical problems which exist in the world with which we must deal presently and with great urgency because we are concentrating too much on constructing needless foundations for what is “manifestly unjust.” Thus Sen’s reasoning in the following paragraph:

I would like to wish good luck to the builders of a transcendently just set of institutions for the whole world, but for those who are ready to concentrate, at least for the moment, on reducing manifest injustices that so severely plague the world, the relevance of a ‘merely’ partial ranking for a theory of justice can actually be rather momentous.³²

This belief, however, is wrongheaded. We can actively pursue what we perceive as justice while at the same time theorizing about the fundamental bases of justice; there is no fault, and indeed great laudability, in this. Of course, we will have to remain open to criticism in the process of pursuing what we perceive as justice in order to further refine and perhaps even eventually completely change what we perceive to be just, and this is precisely what a careful foundationalist approach ensures. By being explicit about our foundations, we may thus expose them to all, including to ourselves, for criticism and re-appraisal; we must, however, at least have a foundation with which to start. The problematic nature of Sen’s assertions is not manifest in his views on such evils as famine; we should not flinch when we have the means and opportunity to help a famine victim, for we can most likely all agree that it is rather reasonable to assume, whichever standards of goodness we are using, that famine is bad. The trouble comes, rather, when Sen makes such assertions as the following:

Partial orderings... can have quite a significant reach; for example, if it is agreed that the status quo in the United States, which does not come anywhere close to universal medical coverage, is distinctly less just than a number of specific alternatives which offer different

³² *Ibid.*, 263.

schemes of coverage for all, then on the grounds of justice we can reject the status quo of non-universal coverage, even if reasons of justice do not fully rank the alternatives that are all superior to the status quo.³³

The fact is, however, that it is not the case that all people agree that non-universal coverage is a bad thing. In fact, the debate over whether health-care *coverage* or *access* is more important is a significant one in the health-economics literature.³⁴ It is thus by no means incontrovertibly clear that universal coverage is the best state of affairs; perhaps universal or near-universal access is. Many people, although perhaps not the majority, furthermore believe that the status quo in U.S. health care is quite good. What do we say to these people? How well will an assumption-based partial ranking of values do us then? Sen would tell us that “If someone has the power to make a difference that he or she can see will reduce injustice in the world, then there is a strong and reasoned argument for doing just that.”³⁵ But which “reasoned argument” is that, and which criteria should we use to discern just states of affairs from unjust ones? Sen tells us that we should be “impartial” and value equality, and that we should also value happiness and freedom. Using these values, we can effectively compare states of affairs and make decisions concerning what is just and unjust. But, again, how exactly are we to do this? My question is not that which Sen answers, namely “Which of these values is more valuable than the others?” but, rather, “How are we even to go about comparing them in the first place?” The problem here, common to all pluralist theories, is one of coherence constraints. If one’s theory is to be pluralist, one must at least have a singular principle that allows us, if not to make them commensurable (an idea which Sen criticizes), at least to show us the relationship they have to each other so that we may meaningfully compare them. It is not clear, furthermore, how Sen could even form his so-touted “partial rankings” of values if he does not have a method or principle whereby it is possible to compare values in the first place. Of course, Sen often promotes the use of “reasoned scrutiny” in order to compare values, but this is hardly an adequate method; this is comparable to my answering “use mathematics” in response to a question concerning how to integrate a function.

As his argument advances, Sen’s pluralism becomes more and more incompatible and, yes, incommensurable. He criticizes the idea of a “single-focus

³³ Ibid., 400.

³⁴ Lynn A. Blewett, et. al. “Monitoring the Uninsured: A State Policy Perspective.” *Journal of Health Politics, Policy and Law*. 29, no. 1 (2004): 107–145; M. Kate Bundorf and Mark V. Pauly. “Is Health Insurance Affordable for the Uninsured?” *Journal of Health Economics*. 25 (2006): 650–673; John C. Goodman, Gerald L. Musgrave, and Devon M. Herrick, *Lives at Risk: Single-Payer National Health Insurance Around the World*. (Lanham: Rowman & Littlefield Publishers, Inc.: 2004); Jonathan Gruber, “Covering the Uninsured in the United States.” *Journal of Economic Literature*. 46, no. 3 (2008): 571–606; Scott E. Harrington, “The Health Insurance Reform Debate.” *The Journal of Risk and Insurance*. 77, no. 1 (2008): 5–38; Paul Starr, *Remedy and Reaction: The Peculiar American Struggle Over Health Care Reform*. (New Haven: Yale University Press: 2011).

³⁵ Ibid., 271.

understanding of freedom” because “freedom as an idea has irreducibly multiple elements.”³⁶ Freedom, as well as indeed almost all concepts, is comprised of multiple elements, but does this pose a problem for foundationalism or monism? I cannot see how. Once again, if we consider *telos*, and specify its exact nature, we will be able to compare the multiple elements of states of affairs by their relative tendencies to achieve that *telos*, no matter whether these elements are split into two or two billion parts. Sen seems to ignore the possibility, for instance, that there is a best form of freedom, one that can be justifiably valued and promoted above all other forms of freedom.

Let us make no mistake; Sen’s theoretical aims are laudable. His emphasis on behavior as against institutions is very important work for all value theorists to consider. His complete disregard for foundationalism and particularly transcendental institutionalism, however, is unfounded. While it is true that many of the rigidities and assumptions inherent in transcendental institutionalist theories ought to be revised, his and others’ views concerning the problematic nature of foundationalist approaches is, indeed, the exact reverse of the true state of affairs. While anti-foundationalists typically charge foundationalists with a disregard for practical matters, with utopianism, and with stiling the progression of ethical thought, there is yet much reason, as I have shown, to state that it is actually the aims of anti-foundationalism that contribute to the impracticality of theoretical ethics in its application to social, legal, political, and economic affairs. While anti-foundationalists indict foundationalists on many counts, it should be seen that it is actually anti-foundationalists who fail to provide us with concrete or at least semi-concrete answers to many pressing matters. When, for instance, is force against an individual justified? When are strikes and lockouts, for instance, justifiable or unjustifiable, and should the state be involved? Is a free market, a mixed market, a command economy, or something altogether different, best? And, most important of all, how would we know? While anti-foundationalist accounts typically do not even come close to giving us an answer, foundationalist accounts show much promise and furthermore provide a sort of check, subject to public criticism, which enables a constant evaluation and re-evaluation of claims. This check is effective, however, only if there is *some* foundation which is, implicitly, *possibly* flawed.

2.2.3 *Karl-Otto Apel’s Rebuttal to the Münchhausen Trilemma and Its Relevance to Value Theory*

In one of his insightful rebuttals to the well-known Münchhausen Trilemma, Karl-Otto Apel explains how discourse theory itself enables us to demonstrate decisively that anti-foundationalist epistemology, and by extension anti-foundationalist ethics, is incoherent and untenable. “From the point of view of transcendental pragmatics,”

³⁶ *Ibid.*, 308.

he states, “the deductive process by which sentences are deduced from sentences, indeed, all ‘axiomatics,’ can only be considered as an objectifiable *means within the context of the argumentative grounding of statements through knowledge-evidence.*”³⁷ In other words, to state that the axioms of logic and mathematics can be only circularly justified or justified via an infinite regression is to miss the salient fact that such axioms cannot be rejected without a necessary appeal to them. Doubt of their *a priori* necessity, then, “is not explainable as a meaningful language game without in principle presupposing at the same time indubitable certainty.”³⁸ This epistemological attitude, as Apel himself demonstrates, has great immediate relevance to ethical theory, for it suggests that, just as it is not possible to cross a certain line of meaningfulness in epistemology, so it is not possible to cross a certain line of meaningfulness in value theory either. We are perhaps again reminded of Richards’s evolutionary account of ethics and its implicitly grammatical basis, as well as the partial basis for my rebuttal of Street’s biological argument. So far, the reader sees only bits and pieces of a framework of value coming together but does not yet see its determinate content or form. The time is not yet right, however, for such an account; it is sufficient for now to have introduced the reader to the general type of method I will employ. Needless to say, the method will be relevantly similar in nature to those cited but will diverge from them in important respects. For now, it is time to leave the background introduction I have laid out and to proceed with the first of several central premises in my argument: that concerning the essence of rationality.

2.3 The Concept of Rationality: Toward a Universal Model

In this section, I use insights from the fields of biological anthropology, social cognitive theory, neuroscience, and philosophy to analyze the concept of rationality. As I demonstrate, prominent conceptions of rationality need to be radically revised. Rationality, I argue, is not solely a deductive and inductive logical phenomenon; it is also very importantly an experiential and perceptual phenomenon. To elucidate this, I introduce the concept of what I call “experiential rationality.” Although other theorists, especially those in phenomenology, have periodically analyzed this aspect of rationality, most ethical theorists in the modern period have either severely underemphasized its immense relevance to ethics or have failed to recognize such relevance altogether. There are important aspects of experiential rationality that serve as justificatory grounds of value theory, but, while highly psychological and phenomenological, they are neither essentially emotive nor hedonistic. Once we have

³⁷ Karl-Otto Apel, “The Question of Grounding: Philosophy and Transcendental Pragmatics of Language.” *Karl-Otto Apel: Selected Essays, Volume Two: Ethics and the Theory of Rationality*. Ed. Eduardo Mendieta. 1996. Atlantic Highlands: Humanities Press, 76.

³⁸ *Ibid.*, 79.

discovered the core, and thus the most general and generic, concepts that comprise rationality, we can, I argue, more clearly delineate and characterize the linguistic and conceptual framework that rationality comprises. As I shall show, a sufficient level of generality can be achieved to make the model applicable to all possible conceptions of rationality, and it can do so without being rendered useless or uninteresting. This model will become, so-to-speak, one of the main premises in my argument for a post-dialectical, aretaic value theory.

2.3.1 Concepts, Conceptions, and the Possibility of a Universal Model

Out of curiosity, I once asked my mother where she thought morality came from. Her response was intriguing. “From the soul,” she said. “It’s from the whole person.” Her response approaches a fundamental truth, one that virtue theorists have long held. No, this truth is not that goodness is a simple, non-natural property, that moral truths are based upon irreducible pluralities, or that the moral self is located in a noumenal reality. This truth, rather, is simple: ethics does, indeed, reside within us. This is not to say, either, that ethics is subjective or relative. On the contrary, ethics seems to be based upon objective truths. These truths, however, can only be found within us, by delving into who and what we are. Macintyre has strongly urged that humans cannot be understood, especially in a moral sense, unless we interpret them *qua* society. While this is true in certain respects, especially with regards to the problematic manner in which reasoning begins, namely upon prejudices, it does not follow from this that humans do not possess fundamentally identical faculties and capacities that we can both empirically model and philosophically analyze and from which we can discover timeless and cultureless facts. How, in fact, would it be possible to study humans *qua* society other than to understand, first, their essential characteristics? It seems that, without this initial understanding, all subsequent analysis of humans *qua* society would be as vacuous as the analysis of A *qua* B. In any case, it is not humans *per se* that ethics is primarily concerned with, although ethical analyses are incidentally concerned with them. It is, rather, persons that we seek to understand, and the properties comprising personhood are supervenient to whatever material embodiment persons might take. Rationality, the essential characteristic of persons, is a complex concept, and we must take great care to accurately define and characterize it. It must be acknowledged, however, that, in order for a theory of rationality to be the most widely applicable and thus the most relevant to as many paradigms as possible, it must be construed and understood as a concept, instead merely as one of many conceptions. In order to do this, we will have to show why a particular reduction is adequate and thus why it is impossible to further reduce or generalize the model. Of course, all models are models of *something*, and so it might be argued, on my presentation, that the model, *qua* model, does not accurately depict reality. This, of course, is true,

and I do not seek to show that such a model accurately *depicts* reality (in other words, that it is showing in full the reality of the state of affairs) but, rather, that it accurately *reflects* reality. In other words, I seek to demonstrate that it does the best job any model can possibly do at describing, in a universally understandable way via maximal inclusiveness of conceptions, that to which all necessarily refer when they utter “rationality.” In order for us to surpass the model and thus understand that to which the model refers, we would have to experience rationality ourselves, and alas we do. Thus, the model and its correspondence (including the myriad personal and cultural conceptions), should not, if we are careful and thorough, be too far removed from the reality of the situation.³⁹ The concept-conception distinction is, of course, a manifestation of the essence-accident distinction, and it is wise not to overlook this. Macintyre’s account of rationality, for instance, seems to present us with an account of rationalities, as if they were completely different concepts, instead of different and, yes, culturally dependent conceptions, the commonalities of which form the singular concept of rationality.⁴⁰ We can understand this relationship as analogous to the Aristotelian linguistic concept of *pros hen* equivocation. Just as the Greek implies, the various uses of “rationality” advance *toward one* concept, unlike the various uses of the words “bat” or “can.” Although, when we, as individuals, think of rationality, we have some conception tied, invariably, to other concepts such as mathematics, architecture, or even love, there is yet much reason to believe that it is tied to some *common* concept such that we understand each other when we utter such sentences as “He is acting angry and irrational” and “The patient is not rational and, thus, cannot consent to surgery.” In observing others’ behavior in relation to the uses of these words, and in acting on these observations in a like manner, it becomes clear that we understand each other in some fundamental, although perhaps very general, way. Such a notion is indeed salient in Chomsky’s theory of generative grammar.

Macintyre is correct in his assessment that human identity exists along a continuum of historical and traditional narrative, but this is somewhat misleading. What Macintyre should assert, perhaps, is that *humans’*, instead of *human*, identity exists along such a continuum. The distinction, here, is vital. What Macintyre’s word usage would seem to imply is that the human *to ti en einai* or essence can only be understood and analyzed within a culture and tradition, since the human essence is inextricably enmeshed with them. Throughout his work, *After Virtue*, however, he seems to promote this interpretation of human essence while simultaneously emphasizing the moral identity of the individual human in relation to society and not, as he might like to suggest, in relation to humans as a class. This pattern of

³⁹ This qualification derives mostly from how scientists typically characterize models, namely as tools for providing a certain degree of predictive and explanatory power. In other words, models, as scientists typically agree, are not strictly equivalent to reality. The phenomenologists’ elaboration of this, of course, is that only a further inquiry into the nature of reality *qua* perception or qualia is sufficient for the elucidation of reality.

⁴⁰ Alasdair Macintyre, *Whose Justice, Which Rationality?* (Notre Dame: University of Notre Dame Press. 1988).

usage, then, points to a very different conclusion than what he implies. While we can understand the class of humans and, indeed, the class of societies as having some respective essence, this essence supervenes, especially saliently in the case of humans, on an individual human's essence such that that individual's essence constitutes an accident of the class of humans. This accident might take the form of race, gender, sexuality, cultural habits, and overall self-conception, yet this accident, and often all of these accidents combined, constitute the essence of a particular, individual human identity. From this perspective, Macintyre's thesis is more intelligible, for we could reasonably expect that, although a supreme moral principle and *summum bonum* might be constituted by the essential characteristics of persons, it could be expected that these characteristics would be expressed in radically different ways across different historical epochs and cultures. To his credit, Macintyre seems to suggest, in his later writings, that such an objective characterization is both possible and necessary.⁴¹

We face, then, the task of setting out characteristics that form that which is, in the words of Aristotle, "special to a human being."⁴² Once again, for the sake of precision and inclusivity, we will not simply address human beings, but persons and agents. Most readers will note that I have already used an Aristotelian distinction, that of the essence-accident distinction, in my arguments. In order to begin any discourse on rationality, it is vital to address this distinction. It is perhaps in their failure to address this distinction at the outset that most theorists in the modern period have set forth definitions which either fail to have complete relevance to ethical theory or which take into consideration only a narrow range of concepts. This distinction, however, is truly indispensable for understanding why the *concept* of rationality is different from the various *conceptions* of rationality and why, indeed, rationality is an essential property. In addition, this distinction is, despite any lingering objections to it either on existentialist or other highly problematic constructionist grounds, both a real and useful one. David Oderberg provides the most recent, comprehensive account of why this is the case in his work, *Real Essentialism*.⁴³ Virtue theorist, Michael Slote, has also provided a comprehensive metaphysical analysis of the distinction in *Metaphysics and Essence* which, although not recent, persists as a reliable testament to the

⁴¹ Specifically, see:

Alasdair Macintyre, *Dependent Rational Animals: Why Human Beings Need the Virtues*, (Chicago: Carus Publishing Company. 1999).

97: "We have . . . identified two crucial respects in which the virtues are indispensable to human flourishing: without developing some range of intellectual and moral virtues we cannot first achieve and then continue in the exercise of practical reasoning; and without having developed some range of those same virtues we cannot adequately care for and educate others so that they first achieve and are then sustained in the exercise of practical reasoning."

166: "It is because and insofar as rational enquiry serves and partly constitutes that common good that it is itself the good that it is."

⁴² Aristotle, *Nicomachean Ethics*, Ed. Roger Crisp. (Cambridge: Cambridge University Press. 2004): 12.

⁴³ David S. Oderberg, *Real Essentialism*, (New York: Routledge. 2007).

soundness of the essence-accident distinction.⁴⁴ Thus, in lieu of making my own metaphysical arguments here, and thus disrupting our main focus on meta-ethics and value theory, I direct the interested reader to these fine works.

2.3.2 “Humaniqueness,” Social Cognitive Theory, and a Neuroscientific Account of Judgment

In 2008 at Harvard University, Hauser presented an account of rationality which he has termed “humaniqueness.” As might be evident from its name, his account picks out that which is neurophysiologically, cognitively, and perceptually unique about humans.⁴⁵ According to Hauser, there are four fundamental characteristics that set humans, or what I will be speaking of as persons, apart from other animals. The first consists in a complex problem-solving capacity: “to apply the same ‘rule’ or solution to one problem to a different and new situation.” This faculty plays a large role in making possible our inductive and deductive capacities, what philosophers commonly refer to as “laws of logic,” including such laws as *modus ponens*, *modus tollens*, *modus ponendo tollens*, and non-contradiction. The second consists in symbolic thought and communication: “the ability to combine and recombine different types of information and knowledge in order to gain new understanding.” This capacity for recombination and synthesis can be thought of as the faculty that enables our discovery and construction of both simple and complex concepts. The third and closely related faculty is that of the capacity for the interpretation and communication of such concepts, what Hauser characterizes as the ability to “create and easily understand symbolic representations of computation and sensory input.” Such a faculty, as would be expected, is highly dependent on our capacity for using the “laws of logic.” Hauser characterizes the fourth faculty as a capacity to “detach modes of thought from raw sensory and perceptual input.” Put more simply, this constitutes our capacity for conceptual abstraction. These capacities, then, constitute a rough yet accurate outline of what rationality is and how it constitutes the essence of persons. We should not, however, stop here in our characterization of rationality, since, although Hauser’s observations are instructive, they tell only one part of the story.

Albert Bandura, one of the most influential psychological theorists of our time, has developed a comprehensive account of agency through years of rigorous and careful empirical research. His theory demonstrates the rich cognitive framework that constitutes agency and, thus, provides a strong case against behaviorism and

⁴⁴ Michael Slote, *Metaphysics and Essence*, (Oxford: Basil Blackwell. 1974).

⁴⁵ Amy Lavoie, “Hauser Presents Theory of ‘Humaniqueness,’” *Harvard Gazette*. 14 Feb. 2008. 13 April 2009. <http://news.harvard.edu/gazette/story/2008/02/hauser-presents-theory-of-humaniqueness>

reductionism. His theory is constructed as a multi-tiered view of consciousness, divided into at least three levels of causally interactive and interdeterministic agency. As Rottschaefer explains, Bandura's theory is one of reciprocal determinism in that "environmental, behavioral, and cognitive factors interact with each other in effecting behavior."⁴⁶ To understand more coherently what Bandura means by "reciprocal determinism," one must first become more familiar with the functions and characteristics of each level of agency. Although Bandura has characterized each aspect of agency *in extensio*, Rottschaefer has labeled them each in a specific way, from the lowest to the highest level of agency. To assure the reader that neither I nor Rottschaefer are sneaking into our explanation any covert normativity via hierarchy, I will first explain what is meant by "lower" and "higher" levels. The relationship between higher and lower levels of agency can be expressed in the following way: A is a higher level of agency than B if A must be characterized in terms of B, and B is a lower level of agency than A if B need not be characterized in terms of A. In addition, higher levels of agency tend to have the capacity to regulate and override the lower levels to a greater extent than the lower levels tend to have the capacity to do so to the higher levels.

The lowest level is base-level agency. Rottschaefer characterizes this level as consisting of "end-governed propensities to perform certain behaviors." These propensities, he explains, "can be either primarily genetically based or primarily learned. They have representative and motivational features, but they are not fully cognitive in the way that beliefs and desires that are propositional attitudes are."⁴⁷ This is the level of agency to which many ethical theorists link first-order desires; it is also the level which is least voluntary. The next highest level Rottschaefer terms behavioral-level agency. This level, he explains, "is composed of beliefs and desires about actions to be performed, and it is these that directly bring about action." This level is thus intricately involved in the formation of second-order desires, or desires concerning first-order desires. At this level, the self is conceived of as a cognitive agent such that the agent acts based on beliefs and desires which are both self- and other-directed. It is at this level of consciousness that Bandura characterizes the agent as a cognitively motivated actor, although this level does not, in itself, exhibit self-evaluation. The next level, that of reflective agency, also concerns beliefs and desires, but is far more complex than the other two levels due to its metalevel motivational structure. This level is characterized not only by its ability to form higher-level beliefs and desires about one's lower-level beliefs and desires, but also by its self-evaluative nature. Beliefs, desires, and ultimately decisions at this level are formed on the basis of some "acquisition, maintenance, and application of self-standards."⁴⁸ Rottschaefer explains the complexity of this level and its relationship to autonomy and voluntary action. He states:

⁴⁶ William A. Rottschaefer, *The Biology and Psychology of Moral Agency*, (Cambridge: Cambridge University Press. 1998): 131.

⁴⁷ *Ibid.*, 19.

⁴⁸ *Ibid.*, 135.

The beliefs and desires of the reflective level among which are moral norms, however these are identified, work directly on the first-level beliefs and desires. Moral action is a feature of the operation of the third-level system, the reflective level. On this model, free actions are those that proceed on the basis of the beliefs and desires of the third-level motivational system. In a free action they determine the beliefs and desires of the behavioral level, as well as the base motivations, that will bring about the action. A relatively maximal condition on freedom would require that the reflective level be operative. A minimal one requires merely that the action performed by base-level or behavioral-level motivational systems be of the sort that a third-level intervention was possible, that is, that they not be compulsive.⁴⁹

Bandura himself terms this phenomenon “reflective self-consciousness.” As he explains, it is this sort of self-consciousness which “enables people to analyze their experiences and to think about their own thought processes.”⁵⁰

This account of voluntary action, viewed as a function of the reflective level’s interaction, in some capacity, with the lower two levels, substantiates Gewirth’s account of agency. Gewirth explains:

A moral judgment is reflective: it does not consist merely in conforming to some accepted practice; it also carries with it an implicit claim, on the part of the person who sets forth the judgment, that he has made it after due consideration and that it is right or correct. The moral judgments made by A, that B ought to do X (where B may or may not be identical with A), is regarded by A as having a good reason in its support, even if this reason consists only in the value of adhering to custom.⁵¹

Bandura’s account also suggests the soundness of Apel’s understanding of rationality, one that is fundamental to discourse ethics and more specifically, I believe, to a foundationalist re-interpretation of Habermasian theory. Apel identifies “reflection” as “the forgotten or disclaimed type, or even method” of what he calls “philosophical rationality,” which he characterizes as one of “transcendental-pragmatic reflection on the normative conditions of its intersubjective validity.”⁵² It is this engagement with the concept of reflection inherent in rationality, I believe, which enables Apel to extend his interpretation even further into the realm of the experiential, as can be seen somewhat in his concept of “hermeneutical discourse.” While his analysis of the experiential is, as we shall see, limited, it is partially through Apel’s framework, supported by Bandurian theory, that we begin to see the emergence of a link between reflection and the experiential.

Rottschaefer establishes yet another level of agency, the self-system, provided by Bandura’s model. This so-called self-system is what Rottschaefer terms “self-referential.” This level, he explains, includes “conceptions of the self and motivations arising from... self-conceptions, in particular, conceptions of the self as a

⁴⁹ Ibid., 19.

⁵⁰ Bandura, *Social Foundations*, 21.

⁵¹ Gewirth, *Reason and Morality*, 14.

⁵² Apel, “Types of Rationality Today: The Continuum of Reason Between Science and Ethics.” *Karl-Otto Apel: Selected Essays, Volume Two: Ethics and the Theory of Rationality*. Ed. Eduardo Mendieta. 1996. Atlantic Highlands: Humanities Press, 143, 149.

moral agent.”⁵³ More specifically, this level is characterized by at least two interrelated subsystems. The first of them, the self-evaluative system, is characterized in terms of the reflective level in that it is “constituted by competencies for forming self-standards and reacting evaluatively to efforts in attaining these standards.” In other words, while the reflective level acquires, maintains, and applies self-standards, the self-evaluative system works deliberately to form and, if need be, to alter these self-standards. The other subsystem is the self-efficacy system. This subsystem is “formed by competencies for self-efficacy judgments.”⁵⁴ This subsystem, Rottschaefer explains, is concerned with judgments of self-efficacy or one’s capacities for certain types of action. As he states: “Self-efficacy judgments. . . reveal both an element of reflexivity about capacities and competencies and an element of self-referentiality. The self involved here is the potential self, in contrast to the ideal self of the self-standards.” This subsystem, and the self-system in general, thus, functions to construct and devise self-standards based on a certain conception of the self or self-image and not, as in the reflective level, simply to act on self-standards that are already present. From this perspective, Rottschaefer explains, “Being moral has become a part of what it means for the person to be the kind of person he or she aspires to be.”⁵⁵ One can see, then, why such psychological accounts of agency are so relevant to virtue ethics, for not only do they reveal the essential elements of practical rationality, but they also reveal that, although standards for action exist, and although we might categorize these, in some sense, as principles or rules, such moral standards are, at the same time, tightly bound up in one’s personal self-conception or, as the virtue theorist might assert, one’s character. As social-psychological theorist and philosopher Nancy Snow has recently argued in her refutation of situationism, character traits that constitute virtues derive from a complex cognitive-affective processing system which is capable of cross-situational consistency and regulation.⁵⁶

Although these social cognitive accounts of rationality and agency do not, by any means, disprove metaphysical determinism, they do pose a challenge to many naïve conceptions of determinism which conceive agency in terms of initial physical causes and resulting mental events. From Bandura’s theory, we can clearly see that, from a deterministic standpoint, agency is a much more complex phenomenon, for several sorts of causation, and not simply one, are at work. From the perspective of emergent mentalism, the theory demonstrates that neurophysiological subvenient events cause supervenient mental events within the self-system and other systems, that the self-system causes events within itself and other systems, and that the self-system and other systems cause neurophysiological events. As Bandura’s theory thus demonstrates, judgment is a key factor among this causal complexity.

⁵³ Rottschaefer, *The Biology and Psychology of Moral Agency*, 19.

⁵⁴ *Ibid.*, 132.

⁵⁵ *Ibid.*, 146.

⁵⁶ Snow, *Virtue as Social Intelligence*.

This account of judgment's role in agency is further substantiated by renowned neuroscientist, Michael Gazzaniga. The brain, he asserts, "acts on its own before we become consciously aware of its actions."⁵⁷ Moral judgment, he explains, is best accounted for by thinking, not in terms of free will, but in terms of what he and other neuroscientists call "free won't." Citing the work of Benjamin Libet on the readiness potential, he explains how we might understand this phenomenon from a neurophysiological perspective. He states:

Libet measured the activity of his subjects' brains using a technique known as event-related potentials (ERPs) while the subjects made a conscious and voluntary hand movement. . . . Libet found that even before 'time *t*,' when the subject first became consciously aware of his decision to make a hand movement, the subject's brain was active—the readiness potential was present. The time between the onset of the readiness potential and the moment of conscious decision-making was about 300 milliseconds. If the readiness potential of the brain begins before we are aware of making the decision to move our hand, it would appear that our brains know our decisions before we become conscious of them. . . . Libet argued that because the time from the onset of the readiness potential to the actual hand movement is about 500 ms, and it takes 50 to 100 ms for the neural signal to travel from the brain to the hand to actually make the hand move, 100 ms are left for the conscious self to either go with the unconscious decision or veto it. That, he said, is where free will comes in—in the veto power.⁵⁸

Judgments, therefore, form an essential part of our voluntary action, including thinking in general. Although this might seem like a truism to some readers, this kind of evidence is necessary for what is to come in later theoretical analysis, and also to satisfy some critics of Gewirth who have had trouble accepting his dictum: "Every agent implicitly makes value judgments—there are no indifferent rational actions."⁵⁹ Indeed, as I have preliminarily suggested, such a statement concerning reflective value judgment lies at the core of discourse ethics as well.

The accounts of Hauser and Gazzaniga, we also see, subsequently reinforce Bandura's theory, especially as regards the manner in which he claims cognitive phenomena arise from neurophysiological activity. As he states: "These advanced neural systems for processing, retaining, and using coded information provide the capacity for the very characteristics that are distinctly human—generative symbolization, forethought, evaluative self-regulation, reflective self-consciousness, and symbolic communication."⁶⁰ It is in precisely these terms that he characterizes higher-level cognitive phenomena, or what I refer to simply as "rationality." This is directly consistent with Hauser and Gazzaniga's respective accounts.

⁵⁷ Michael S. Gazzaniga, *The Ethical Brain*, (New York: Dana Press. 2005): 93.

⁵⁸ *Ibid.*, 92.

⁵⁹ Gewirth, *Reason and Morality*, 40.

⁶⁰ Bandura, *Social Foundations*, 21.

2.3.3 *Inadequacies and Limitations of Competing Accounts*

Now that we have explored at least some of what is understood empirically about rationality, how can we package this together with a philosophical method to make it coherent and relevant to ethics? It is at this point that the concept-conception distinction becomes vital, for if one is going to develop an accurate account of normativity based upon rationality, one better demonstrate why it is superior to all of the other definitions and characterizations of rationality that philosophers and ethical theorists have already proposed. In order to achieve this, as I stated earlier, we must provide, not a conception of rationality, but the *concept* of rationality to which everyone implicitly refers when forming any *conception* of rationality. Furthermore, we must also explain how such a concept, as general as it might be, could possibly be relevant to ethics. As I will demonstrate, both of these tasks are surmountable.

Within the field of ethics, theorists have established various definitions of rationality, most of which have seemed to be tailored to suit the needs and contingencies of their specific theories but which therefore exclude what I take to be its most fundamental elements. Kant famously set forth the distinction between autonomy and heteronomy to deal with the complex phenomenon of agency. Autonomy, he asserted, has its characteristic feature of freedom by virtue of its noumenal nature. Free will, on Kant's account, is a vital feature of right action and is, therefore, characterized by its adherence to the categorical imperative. Heteronomy, on the other hand, is characterized by its adherence to some hypothetical imperative, whether it be custom, sanction, or some other motive.⁶¹ Heteronomy, for Kant, is characterized, thus, by its lack of freedom. In her work, *The Sources of Normativity*, Korsgaard presents an account of agency and rationality along the same lines as Kant.⁶² Unfortunately, however, these sorts of Kantian and Neo-Kantian accounts of agency fail. As Fritz McDonald explains: "One troubling aspect of such a narrow definition of agency is that it would rule out the possibility of irrational or immoral acts, insofar as such agency is not based on universal principles of rationality and morality. On Korsgaard's account, the only way to be an agent is to act on rational principles, including the categorical imperative. If this is so, it is difficult to see how immorality or irrationality is even possible."⁶³ In other words, since one cannot technically act voluntarily if one is acting in accordance with hypothetical imperatives, then one cannot be morally responsible for any adherence to such imperatives. It follows from this that there can be no immoral actions. This, however, is both highly implausible and self-defeating. Not only have

⁶¹ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, (Cambridge: Cambridge University Press. 1997).

⁶² Christine Korsgaard, *The Sources of Normativity*, (Cambridge: Cambridge University Press. 1996).

⁶³ Fritz J. McDonald, "Agency and Responsibility," *The Journal of Value Inquiry*. 44, no. 2, (2010): 199–207.

Hauser, Gazzaniga, Bandura, Lawrence Kohlberg, and a host of others demonstrated that such an account of agency does not accord with the evidence, but the soundness of such an account also seems to hinge on the soundness of hotly contested and doubtful metaphysical premises.

Bernard Gert, a self-proclaimed Hobbesian, has set forth what he calls a hybrid theory of rationality, constructed to accommodate “common morality,” or, as he more precisely outlines it, “an informal public system applying to all rational persons, governing behavior that affects others, and includes what are commonly known as the moral rules, ideals, and virtues and has the lessening of evil or harm as its goal.”⁶⁴ While this characterization of morality seems innocent enough, one can only begin to consider accepting this framework if one accepts Gert’s conception of irrationality, which he sets out first. This conception, tailored to meet the requirements of his theory, is unfortunately narrow and question-begging. An action, he states, “is irrational in the basic sense if and only if it is an intentional action of a person with sufficient knowledge and intelligence to be fully informed about that action, and who, if fully informed, (1) would believe that the action involves significantly increased risk of his suffering death, pain, loss of ability, loss of freedom, or loss of pleasure, and (2) would not have an adequate reason for the action. All other intentional actions are rational.”⁶⁵ It is from this definition that he eventually derives his “ten justified general moral rules.”⁶⁶ In the first place, it does not follow simply from the fact that the way most or even all people use the word “irrational” *includes* this sense that people use “irrational” *exclusively* in this sense with regards to other persons. Again, as in Korsgaard’s theory, we see only an emphasis upon judgment, as opposed to experience. We do not, in other words, see in this definition any mention of the perceptual make-up of rationality or the qualia that are entailed by it. The only concepts in his definition that come close to accommodating this are freedom and pleasure, and even those are related to rationality merely via their instrumentality for “action.” His definition, in addition, contains unjustified normative assumptions. Gert, of course, would tell us that these are not simply assumptions but that the values invoked are universal from the perspective of rational self-interest. Even so, it does not follow from this that, in order to be an agent, one must *logically* value these things or cease to be an agent. Can we not think of a highly advanced artificial intelligence, which we could conceivably (although perhaps controversially) call a person, that might act voluntarily but which might not value pleasure, or a person who might want to have certain injuries or impairments out of some eccentric desire? Gert might yet insist, on Hobbesian grounds, that since all agents would accept this account for themselves, it needs no further validation. This approach, however, is what is troubling to all contractarian or contractarian-like theories, including those of such opposed

⁶⁴ Bernard Gert, *Morality: Its Nature and Justification*. (Oxford: Oxford University Press. 1998): 26.

⁶⁵ *Ibid.*, 39.

⁶⁶ *Ibid.*, 216.

positions as John Rawls and Robert Nozick.⁶⁷ Since all of these theories depend on agreement, they have no ultimate justification beyond what everyone simply agrees to out of fear, voluntary association, need, or self-interest, and thus the formal-procedural element is present while the contentual-substantive element is lacking. Gert's further response might be that there exists no such further "ultimate justification" and that systems of ethics among societies can only be formed in this way. Not only is such an ultimate justification possible, as Gewirth and the discourse theorists have shown (Habermas implicitly and Apel and some others explicitly) and as I will elaborate in further sections, but justification via agreement alone is actually fallacious along the same lines as the fallacy of appeal to the majority. A final appeal to "common sense" might be in order or even, perhaps, an appeal to stop "splitting hairs." The rigorous analytical approach that I wish to take, I contend, is not merely an act of "splitting hairs." Common sense, while seemingly harmless, ought more properly to go by the name of "common prejudice," which is, more often than not, what such common sense really is. If we are to be exhaustive about our analysis, we must strive not to let any normative presuppositions seep into our theories and render them question-begging from the outset.

David Phillips has also analyzed Gert's definition of rationality and irrationality and has criticized the account on the grounds that its hybrid nature renders it problematic.⁶⁸ Due to its hybrid nature, he asserts, Gert's account ends up trying to accommodate too much. As he states: "Gert allows for subjective discretion, by saying that a person may rationally choose to act on what is not objectively the strongest reason. Gert is also concerned to emphasize more generally the idea that his theory makes rationality less constraining than certain other theories do. There are many situations when rationality does not dictate a single course of action, but instead when a wide range of possible courses of action are rationally permissible." This, he claims, causes problems for his account of rationality, namely those which are common to all hybrid theories of rationality. Hybrid theories, he states, meet two main challenges: "First," he asks, "if considerations of the good of others are sufficiently powerful that they rationally may outweigh considerations of an agent's good, how can it also be rationally permissible to give them no weight whatsoever?" Gert's theory seems not to provide a satisfactory answer to this very pressing question. There is also the second challenge: "How can a hybrid view possibly be the right way to accommodate simultaneously what is right about egoism and what is right about person-neutral accounts of rationality?" This challenge is reminiscent of a similar attempt on the part of Peter Singer at a *via media* between two disparate accounts of ethics.⁶⁹ Here, as in Gert's account, we find more questions raised than answered.

⁶⁷ John Rawls, *A Theory of Justice*, (Cambridge: Harvard University Press. 1971); Robert Nozick, *Anarchy, State, and Utopia*, (New York: Basic Books, Inc. 1974).

⁶⁸ David Phillips, "Gert, Sidgwick, and Hybrid Theories of Rationality," *The Journal of Value Inquiry*. 35, no. 4, (2001): 439–448.

⁶⁹ Peter Singer, "The Triviality of the Debate Over 'Is-Ought' and the Definition of 'Moral,'" *American Philosophical Quarterly*. 10, no. 1, (1973): 51–56.

While such theories based on Kantianism, contractarianism, and various others might base themselves on seriously problematic accounts of rationality, it might seem as though we could at least accept the accounts of rationality set forth in various versions of game theory, expected-utility theory, and other incarnations of rational-choice theory. Surely, these theories are based on conceptions that are sensible and straight-forward, right? Not exactly. David Gauthier, one of the most renowned proponents of game-theoretical ethics, has put forth a conception of rational decision-making which he claims solves many of the problems concerning the ambiguity of short-term versus long-term preferences inherent in other models.⁷⁰ The same problems that pervade all such theories, however, persist in Gauthier's account.

The main problem inherent in these theories is one of normative presupposition. For a clear example, let us take the classic Prisoner's Dilemma. According to game-theoretical models, prisoner A should defect instead of cooperate, since she will not have to serve as much time in prison if she confesses; she would, thus, be "better off" making this decision. There are, however, three glaring problems with this conclusion. The first, yet again, is of normative presupposition. It is certainly the case that individuals with certain dispositions or desires not to be imprisoned for a specific reason will choose to confess and, therefore, to go free. It does not, however, follow that all or even most individuals in such a situation will make such a decision or that such a decision is, as the claim goes, "rational." If, for example, I truly believe that I have done something horrible and that I deserve to be punished, I might feel guilty enough or think it in my best interest, or even think it heroic, to do whatever it takes to have myself imprisoned for my offense, even if that includes, in this situation, not confessing. Suppose, also, that I have undergone an intense degree of deliberation to come to my conclusion and that, based on the things that I value in life, I have made the decision to be imprisoned. Thus, we can see the problem more clearly. Rational-choice theory's general account of instrumental rationality presupposes a *specific* uniform value system for all agents predicated on self-interest or prudence. On this account, certain values are rational, while others are not. While this might be plausible in some sense, game-theoretical frameworks make the gross assumption that individuals in such a scenario should act according to certain assumed, philosophically unvetted norms. If we accept this framework, we will simply have to accept that we are irrational if we do not value what game theorists tell us to value. My critique encompasses a set of problems Amartya Sen has also focused upon in his call for greater emphasis on what he calls the process aspect of freedom.⁷¹

The second issue here is, accepting for the sake of argument that all agents act on self-interest, that all such self-interest will look the same or even relevantly similar in this dilemma. Whatever my choice for defecting in the Prisoner's Dilemma, I

⁷⁰ David Gauthier, *Morals by Agreement*, (Oxford: Oxford University Press. 1986).

⁷¹ Amartya Sen, *Rationality and Freedom*, (Cambridge: Harvard University Press. 2002); *The Idea of Justice*.

need not do it in some readily predictable, determinate way; much is dependent on my psychology. I might, for example, choose not to confess and, thus, to go to prison because I want my mother or my community or my friends to be proud of my decision to do the right thing, as they and I see it, thus appealing to my own understanding of my own self-interest. Self-interest, and even altruism, for agents is not so neat and tidy as game theorists would have it seem. To understand an agent's behavior in such situations, one must know an agent's psychological dispositions and attitudes, and such information cannot be inferred from game-theoretic scenarios alone. In addition to its conception of rationality, therefore, game theory's conception of self-interest seems to be both normatively presumptuous and poorly defined. Nozick, at least, has addressed such an issue with his addition of what he terms "evidential expected utility" and "symbolic utility."⁷² While these additional variables contribute substantively to the decision-value formula and, thus, widen the scope of what can be considered rational, his account fails to demonstrate how, with these additional variables and a subsequent potential for degrees-of-freedom problems, decision theory retains strong predictive, let alone normative, power.

The third glaring problem with game theory and rational-choice theory more generally is just this: that it is prone to temptations to derive a normative system solely from human behavior, with little regard to the semantic or logical analysis of ethical concepts. In other words, it derives norms about what agents *should* do from what agents *would* do, when in reality, such a derivation of "should" from "would" is logically illicit. We have, perhaps, in both game theory and expected-utility theory, further glimpses at the legacy of behaviorism in ethics. This, I believe, is justifiable to infer, since, just as behaviorism does, these theories of rationality treat agents as base-level or behavioral-level organisms such that, if presented with some stimulus, agents' behavior can be easily predicted via some simple algorithm or formula. These theories, like behaviorism, pretend that the other levels of agency, with their metalevel motivational and reciprocal deterministic structure, do not exist.

A classic objection to these arguments is that "rationality" does not denote or prescribe specific preferences or value systems at all and, instead, that it merely denotes a process whereby agents reason to attain the objects of particular preferences that they do have, whatever these may be. This, however, actually undermines game-theoretical models of ethics, for, in such a scenario as the prisoners' dilemma, if we assume that the prisoners have highly non-standard values, it becomes increasingly difficult, and perhaps even futile, to attempt to predict their behavior and, especially, to normatively prescribe their behavior. An objection to this would be that game-theoretical models take into account differing value or preference systems via increasingly complex prisoners-dilemma and other similar scenarios that take into account multiple variables of preference and behavior; standard statistical and set-theoretical methods, in other words, deal with this

⁷² Robert Nozick, *The Nature of Rationality*. (Princeton: Princeton University Press, 1993).

issue quite straightforwardly. Nozick, specifically, might object that it is possible to predict such complex behavior by subjecting the decision-value formula to some sort of neural-network-like feedback system. While possible, one must still, on his account, assign weights to evidential expected utility, causal expected utility, and symbolic expected utility respectively in some way that will inevitably be biased. It is partially because of this that one must also be skeptical of claims that game-theoretical or rational-theoretical models do not attempt to design or prescribe ethical paradigms. Gauthier's theory is one obvious example, but consider, also, the family resemblance in such accounts as Thomas Nagel's and Rawls's, both of which ask us to make certain normative presuppositions based on hypothetical agreement. Game-theoretical models run into trouble when they attempt to be normative: simple agreement or cooperation is not justification. This is a criticism with which Gewirth and Habermas charge Rawls as well.⁷³

Expected-utility theory is not immune to these sorts of defects either. This theory, most salient in the context of microeconomics but widely influential in other fields as expected-value theory, multiplies probabilities of events by respectively assigned values associated with their outcomes and then sums them in order, ultimately, to guide rational decision-making. As it stands, utility in this theory is a vacuous concept without some coherence constraints dictating the "correct" procedure for weighing differently valued goods and for assigning utility values, and such coherence constraints are inevitably normatively presumptuous. In her work, *The Authority of Reason*, Jean Hampton likewise points to this problem and its relationship to the covert normative assumptions within both game theory and expected-utility theory. By examining this relationship, she claims, we can clarify the nature of these normative assumptions. Within both theories, she explains, "We need to answer the question, 'How should an agent reason so as to attain an end, when she has many ends at that time which she also wants to attain?'" "The standard instrumentalist answer," she explains, "is that the agent should figure out which end is 'most important'—via some kind of 'consideration' process—and then reason instrumentally with respect to that remaining end. . ."⁷⁴ This process, as in economic theory, is represented as a conflict between opportunity costs that must be resolved through some ambiguous "cost-benefit analysis." Hampton narrows in on the precise nature of this problem of ambiguity and coherence constraints by examining the characterization of instrumental rationality as set forth in these theories. "Anyone interested in defining instrumental rationality," she asserts, "is going to have to take a normative stand on what these coherence constraints are—a stand that. . . simultaneously involves taking a stand on the structure of a rational

⁷³ Alan Gewirth, *The Community of Rights*, (Chicago: The University of Chicago Press. 1996); Jürgen Habermas, "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism," *Habermas and Rawls: Disputing the Political*. Ed. James Gordon Finlayson and Fabian Freyenhagen. (New York: Routledge, 2011).

⁷⁴ Jean E. Hampton, *The Authority of Reason*, (Cambridge: Cambridge University Press. 1998): 170.

agent's good."⁷⁵ This normative stand, thus, constitutes a judgment of "what counts as a coherent structure of an agent's preferences," what constitutes instrumental reason, and what constitutes a good end. Game theory and expected-utility theory, therefore, while seemingly innocent and reliable frameworks, are loaded with unjustified normative assumptions. The final asylum for these theories might be an appeal to internal reasoning as a sound framework for forming ethical judgments but, as we shall see in Sect. 3.2 on developing a method of justification, internal-reasoning models also fail. Thus, while game theorists and expected-utility theorists make claims both to sound ethical and behavioral models, these behavioral models are found wanting, and their ethical models are found to be completely unconvincing. One of the main reasons why rational-choice models of normativity are inherently flawed is the flawed premise upon which they all invariably begin, namely that the assignment of a word that has some connection to a favored value of the person undertaking the analysis to the formal process being explicated thereby makes that process legitimate or otherwise valuable. In other words, simply because rational-choice theorists assign the word "rationality" to the variable under investigation does not thereby make that variable valuable. Yet this is what every rational-choice account of ethical theory covertly proposes. Since rationality, in rational-choice theory, is merely a dichotomous variable (e.g.: "rational" and "irrational"), and since it is explicitly a *descriptive* theory of behavior, we need not use these terms. After all, all we are after in rational-choice theory is a description of different types of behavior for the purposes of predicting behavior, are we not? We may then call rational-choice theory "typical-choice theory" instead, separating the variable of typicality into "typical" and "atypical," or we may merely use some arbitrary variable ϕ and separate it into ϕ_A and ϕ_B . Once we strip the theory to its bare bones and take away the normatively charged term "rationality," we see more clearly that there is absolutely no basis for building an ethical or otherwise normative theory via these simple behavioral distinctions of "typical" and "atypical" or ϕ_A and ϕ_B . Characterizations of rationality using these distinctions are thus very limited and even potentially dangerous if employed normatively. While Nozick's fine additions to decision theory thus contribute much to what rational action actually entails, his account remains trapped within the limited paradigm of rational choice (or should we say logical choice?), which emphasizes merely the phenomenon of what I will in further pages be calling "logical judgment;" once again, the experiential elements of rationality are ignored. As such, these additions of evidential expected utility and symbolic utility are merely normatively question-begging, and the norms Nozick goes on to establish based in part upon them fail to adequately answer the important underlying question of whether this thing he is calling "rationality" is, itself, valuable.

⁷⁵ Ibid., 172.

Apel too, because of his discourse-theoretical project, has been impelled to make a case against normative accounts of game theory, rational-choice theory, and expected-utility theory. For Apel, such accounts are inadequate because, in thinking solely in terms of them, “nous manquerions l’essentiel et tomberions dans des difficultés insurmontables si nous tentions de comprendre la *rationalité communicationnelle* à partir du point de vue privilégié par la théorie des *jeux stratégiques*, comme cela est suggéré dans la plupart des approches contemporaines.” This, he states, is because, “dans ce cas en effet, la communication par le langage ne serait rien d’autre qu’une *influence réciproque* des agents par les moyens du langage au service de leurs fins *présupposées*.”⁷⁶ This picture, he is correct to imply, cannot be right, for such an account of human action unduly compartmentalizes human motivation. Again, the models of social cognitive theory we have explored fundamentally conflict with such an interpretation. What Apel believes we have reason to accept, instead, is a conception of rationality embodied by a “self-reflexive discursive *logos*,” which serves, once again, as the foundation of his transcendental-pragmatic, dialectically necessary, dialogical approach to ethics.⁷⁷ While suggestive of an incorporation of experiential elements of rationality, his theory too seems to fall back on judgment-focused conceptions, this of course being expected of a theory founded upon the concept of argumentation, as all discourse theories are. Thus, even in his assertion that “*philosophical rationality is not identical. . . with the unreflexive logical-mathematical consistency of axiomatic systems*,” he adds that it is instead identical to “*the consistency of the reflexive self-redemption. . . of argumentative reason*,” which, although a promising jumping-off point for a phenomenological or otherwise experiential account of rationality, proceeds no further than merely toward a judgment-focused account, but this time merely with respect to the phenomenon of argumentation itself.⁷⁸ While Apel contends that the transcendental-pragmatic preconditions of argumentation are the ultimate grounding of ethics and, for that matter, any and all judgment, I believe there is, in addition to argumentation and axiomatic reasoning, an equally

⁷⁶ Apel, *La Réponse de L’Ethique de la Discussion au Défi Moral de la Situation Humaine Comme Telle et Spécialement Aujourd’hui*. (Louvain: Editions de l’Institut Supérieur de Philosophie Louvain-la-Neuve, 2001): 32.

We would miss the essential part and would fall into insurmountable difficulties if we attempted to understand communicative rationality from the privileged point of view of *game theory*, as is suggested by most contemporary approaches. . . In this case in effect, communication by language would be nothing other than a *reciprocal influence* of agents by means of language in the service of their presupposed ends.

⁷⁷ Apel, “The Challenge of a Totalizing Critique of Reason and the Program of a Philosophical Theory of Rationality Types.” *Karl-Otto Apel: Selected Essays, Volume Two: Ethics and the Theory of Rationality*. Ed. Eduardo Mendieta. (Atlantic Highlands: Humanities Press, 1996): 258.

⁷⁸ Apel, “Limits of Discourse Ethics? An Attempt at a Provisional Assessment.” *Karl-Otto Apel: Selected Essays, Volume Two: Ethics and the Theory of Rationality*. Ed. Eduardo Mendieta. (Atlantic Highlands: Humanities Press, 1996): 209.

fundamental basis for ethics in both the pre-linguistic and post-linguistic *experiential* that must not be overlooked.

2.3.4 *The Necessary Methodological Preconditions of Universal Applicability*

One could continue for hundreds of pages in examining problems in various theorists' characterizations of rationality, but the few preceding critiques, I believe, have at least served my point, namely that conceptions of rationality that limit the scope of their application merely to the phenomenon of judgment and which do not rest upon a self-conscious effort on the part of the theorist to develop a universally applicable model of rationality ultimately fail. This failure rests on two basic, commonly overlooked criteria to which any definition of rationality must conform in order to be universally applicable: (1) that it reflect common usage, that is every possible usage, of the term and (2) that it at least be consistent with the predictions and explanations provided by scientific models.⁷⁹ It must, therefore, be a definition of rationality, namely of the concept itself, that is sufficiently general or generic such that a usage of the word outside the parameters of the definition, say, in reference to a chair, the sun, or a brick, would reflect a misunderstanding of the language. A telltale sign of an inadequate definition of rationality, therefore, is the enduring possibility of such statements as, "I don't use 'rationality' that way" or "That is not what I think of when I think of rationality." A definition of the concept of rationality, therefore, must be, in some sense, transcendently presupposed by all other possible definitions; in this sense, it must truly be *pros hen* equivocal. I believe the oversight of such preconditions is one major source of theoretical error in many theories such as those of Kant, Gert, Macintyre, Gauthier, Rawls, and Korsgaard, for in response to all of their accounts of ethical theory, one might overturn their entire structures simply by asking "But why should I accept your definition of rationality?" Virtue theorists, I believe, often develop more accurate accounts of rationality, for they seem to appreciate the complexities and subtleties inherent in its psychological and phenomenological structure. In fact, in noticing that virtue theorists tend to appeal somewhat more strongly to inner mental experience via the analysis of character states entailed and necessitated by the virtue paradigm, and in likewise noticing that non-virtue-theoretical paradigms tend to emphasize the rationality of judgment via the analysis of principles and rules entailed and necessitated by non-virtue paradigms, one might, with some justification, choose to characterize virtue as more explicitly experiential and principle as more explicitly axiomatic. Such a characterization is obviously non-holistic, but while not wholly sufficient on its own, the characterization at least suggests the possibility that, if as truly inadequate as I claim it is, then judgment-focused

⁷⁹ Again, this is not to say that it should be reducible to these models.

conceptions of rationality, unified with experiential conceptions of rationality, might result in a clearer understanding of the relationships between the aretaic and the dialectical-deontic.

Gewirth's account of rationality, as is again expected of a dialectical theory, strongly emphasizes the phenomenon of logical judgment while largely ignoring the relevance of the experiential aspects of rationality. Nevertheless, his attempt at generality and thus universal applicability is highly effective. His characterization, in a mere sentence, is as follows: "I use 'reason' in a strict sense as comprising only the canons of deductive and inductive logic, including among the latter its bases in particular sense perceptions."⁸⁰ The mention of sense perceptions here is crucial. "Rationality," in this definition, is not presupposed by some covert norms. Indeed, who could meaningfully deny that rationality has *something* to do with sense perception? Unfortunately, in none of his works does he present an elaboration on such "sense perceptions" or their relevance to ethics or, most importantly, *value*. Gewirth, too, misses a crucial piece of the puzzle.

2.3.5 *The Model Explicated and Analyzed*

This, then, leads to my own characterization of the phenomenon of rationality. Sense perception, it seems, is important. This, however, is not enough, as many non-human animals also have what could be called sense perception. There are, then, some other vital components. As we saw from Hauser's biological-anthropological model, "humaniqueness" tends to revolve around abilities for conceptual synthesis and conceptual abstraction, both of which are necessitated by practical or theoretical judgment according to laws of logic, what we may refer to as logical judgment. We can see, from this, a very common, prevalent conception of rationality that revolves around inductive and deductive logic. This is what many people think of when they think of rationality, but this is clearly not the case for all people. We might consider, for example, that a hike in the wilderness, or a visit to an art gallery, or some other aesthetic experience often has a significant impact on a person's rationality. For instance, these experiences often promote self-reflection, self-awareness, and self-understanding. Conversely, we might also consider that computer software uses logic. It is at this point that criterion (1) becomes vitally important. Although "rationality," in some sense, can be meaningfully applied to such a process, its use in such contexts makes implicit reference to an aspect of rationality that has been extended to other systems subject to agentive manipulation. To clarify, although we might conceivably meaningfully utter the statement "Software follows rational processes," this use of "rational" makes implicit reference to a capacity that persons possess whereby they can understand such processes in the first place. Thus, we see that the idea of rationality can indeed be applied to

⁸⁰ Gewirth, *Reason and Morality*, 22.

non-agentive entities but, importantly, that we can do so only via an extension of qualities that we possess and that we have exported, so-to-speak, to these other entities. Software's use of logic, for instance, is possible because agents have exported the use of logic to this system. Software's use of "rational processes," construed as "logical processes," then, is indeed traceable to the concept of rationality, but only via its implicit reference to an agentive activity. Thus, there is reason to conclude that all uses of "rationality" do indeed trace their respective meanings from agentive activity.

This, then, highlights a crucial feature of rationality: it does, indeed, encompass a capacity for logic, but it is logical *judgment*, and not simply logic itself, that is what is at issue. Logical judgment requires the capacities for conceptual synthesis and conceptual abstraction that Hauser has put forth, since conceptual synthesis and conceptual abstraction are necessary for processing information in a way that makes our thinking coherent to us. Thinking that is coherent to us, "coherent" meaning information processing by the use of these faculties of conceptual synthesis and conceptual abstraction, is necessary and sufficient for any logical judgment, no matter how simple or complex. Bandura's research on social learning theory substantiates such a claim. As he affirms: "In the social cognitive view, environmental influences affect behavior through a symbolization process. That is, transitory occurrences have lasting effects because the information they convey is processed and transformed into symbols."⁸¹ This is indeed a process of abstraction, and it reinforces Hauser's own emphasis on symbolic communication. Bandura's model likewise substantiates the centrality of conceptual synthesis in agency. This is apparent not only in his more general elucidation of the interdeterministic nature of cognitive processes but also, once again, in his account of social learning theory and, in particular, in his exploration of the concepts of creativity and originality. As he states:

Originality largely involves synthesizing experiences into new ways of thinking and doing things. When exposed to models who differ in their styles of thinking and behavior, observers rarely pattern their behavior exclusively after a single source, nor do they adopt all the attributes even of preferred models. Rather, observers combine various aspects of different models into new amalgams that differ from the individual sources.⁸²

Through its essential role in cognitive activity, therefore, we see that conceptual synthesis and conceptual abstraction do indeed constitute necessary and sufficient conditions of logical judgment. Thus, we can form the following bi-conditional statement:

$$\text{HasLogicalJudgment}(\Phi) \leftrightarrow (\text{HasConceptualSynthesis}(\Phi) \wedge \text{HasConceptualAbstraction}(\Phi))$$

⁸¹ Bandura, *Social Foundations*, 110.

⁸² *Ibid.*, 104.

At this point, however, one might make the objection that logical judgment, that is, judgment according to the laws of logic, is not a necessary condition of conceptual synthesis and conceptual abstraction. Let us think of some likely counterexample. Let us suppose that I am an eccentric painter who paints according to the first sentiments that come to her; the process is quixotic, emotional, and chaotic. It might even be said to be irrational, but is it? Let us consider the fact that, in order for a painter to even approach a canvass and make the *voluntary* decision to paint, she must at least form some practical syllogism or at least follow some hypothetical imperative such as:

1. Approaching and obtaining a canvass and paint are necessary conditions of my painting.
2. I want to paint.
3. Therefore, I ought to approach and obtain my canvass and paint.

After the painter approaches the canvass and has her paint, one might retort, the process is all irrational from there. Let us not, however, be too hasty. In the process of painting, do I not need the laws of *modus ponendo tollens* and non-contradiction to conceive of even the most quixotic ideas? If this were not the case, it would be a bizarre case indeed, for the artist would not be painting using any judgment whatsoever, meaning that her painting would be some non-voluntary action, perhaps some form of compulsion, although even during compulsive episodes, people generally have awareness of their actions, perceive them in some way, and thus make some perceptual judgment according to some law or laws of logic. Snow has also cited important research on goal-dependent automaticity and the cognitive-affective processing system which supports these claims.⁸³ Our objector, however, might be insistent: “Can we not,” she might ask, “imagine a mind in which concepts come about and combine in all sorts of ways but for which this process does not necessitate any sort of judgment whatsoever?” The answer, quite simply, is no: we literally cannot imagine such a scenario. Think about this closely: if there existed such an entity for which syntheses of concepts occurred, this would not be anything like what is meant by conceptual synthesis and conceptual abstraction; instead, this is far closer to data synthesis and permutation. This is not, in other words, what Hauser had in mind when he set out to describe the synthetic and abstractive processes entailed by humaniqueness. The characterization of data synthesis and permutation as conceptual synthesis and conceptual abstraction, we must see, is yet another such extension of agentic behavior. Of course, what can be termed “data transfer” occurs in nature all the time, for instance via light, but the synthesis of such data cannot then be said to be abstracted, nor can it be said even to be a concept, for entailed by the meaning of “concept” is some sort of existence within a mind that at least minimally judges it, for instance, by at the very least acknowledging its existence. Thus, we arrive at the core of what might be called consciousness and, as this is not meant to be a metaphysical analysis of the mind, I will end

⁸³ Snow, *Virtue as Social Intelligence*.

my reply here. I am content with omitting such a metaphysical analysis of the mind because, strictly speaking, it is not necessary for the acceptance of my reply. All that is necessary to accept the reply is (1) the acknowledgment that *data synthesis* cannot be considered *conceptual synthesis* unless such a process takes place in the mind of an entity that at the very least acknowledges the existence of such data or concepts and (2) the acknowledgment that all usages of “conceptual synthesis” with regards to non-agentive processes ultimately derive their meaning from the concept of agentive activity, where “agent” refers to the entity cited in (1).⁸⁴ Again, this is because conceptual synthesis and conceptual abstraction imply a capacity for picking out different concepts and organizing them according to some schema, and such a schema can be organized and constituted only by *judgment*, if even in a very rough sense implying subconscious discrimination between concepts. Even at such a primordially subconscious level, however, as we have briefly assayed, social cognitive theory accounts for the inherent link between conceptual synthesis, conceptual abstraction, and logical judgment. Thus, conceptual synthesis and conceptual abstraction do indeed necessitate logical judgment because even the mere acknowledgment of the occurrence or otherwise the existence of such a process is a form of judgment and, indeed, a form of logical judgment necessitating, at the very least, the law of identity.

Conversely, one might make the objection that conceptual synthesis and conceptual abstraction are not necessary conditions of logical judgment. Let us explore, then, the minimal requirements for something to be considered a logical judgment. A fitting minimal example is that of my flicking a small, balled-up piece of paper across a table. We must keep in mind, yet again, that this is a *voluntary* action; I am not doing this impulsively or otherwise non-voluntarily. When we refer to impulsiveness, of course, we refer to base-level or behavioral-level activity, as Bandura indicates, without any interaction between them and the reflective level. What, then, are the preconditions, that is, the necessary conditions, of my flicking this piece of paper across the table? I must first be minimally aware of my environment such that I can infer that, by cause and effect, if I act on this piece of paper with my hand in a certain way, then it will probably move in a certain way. I must, therefore, access my memory of past observations of relevantly similar scenarios in which such cause-and-effect scenarios occurred. I must then use this information and apply it to this scenario in a relevantly similar manner. This process exhibits a minimal level of conceptual abstraction in the form of the re-application of past perceptual information to this new scenario and a minimal level of conceptual synthesis in the form of the analysis of inductive or deductive relationships and their application to a new scenario containing new information. The mind normally accomplishes all of this very quickly. One further objection might be raised. One might assert, “I do not need to abstract at all in order to make logical judgments. What if I have never experienced anything like a table, a balled-up piece of paper, or (bizarrely) a hand? What if all of this is genuinely new to me?” We should be skeptical of such remarks.

⁸⁴ Note that this is similar to our example “Software follows logical processes.”

How genuinely new could such an experience be? Except for the idea that there are people who are unfamiliar with what hands are, it is conceivable that someone might come from a culture where there are no tables and no balled-up pieces of paper. However, all people in all cultures are aware that objects exist and that objects, whatever they might happen to be, have certain determinate, expected parameters for their behavior within the context of empirical reality. One at least has familiarity with some type of object and, therefore, one can infer some general sort of behavior of the entity *as an object*, even if she is unfamiliar with what kind of object it is. As Bandura affirms:

Imaginal representations are abstractions of events, rather than simply mental pictures of past observances. As a result of repeated exposure to modeled events, observers extract distinctive features and form composite, enduring images of the behavior patterns. Activities are, of course, rarely performed in exactly the same way on repeated occasions. No two tennies serves are identical; no two apples are the very same. Observers have to construct a general conception that encompasses essential aspects from specific instances that vary around a basic pattern.⁸⁵

These “essential aspects,” we must admit, constitute both forms of abstraction and products of conceptual synthesis.

Still, the recalcitrant skeptic might retort, “But infants learn things as they grow. Their expectations of reality do not come from initial abstractions and conceptual syntheses.” Enough scientific evidence exists, however, to demonstrate the flaws in this remark. Human neonates are not agents, although they generally have the capacity to become agents. Eventually, they develop the ability for logical judgment and, thus, conceptual abstraction and conceptual synthesis through their exposure to and cognitive interaction with the environment; these faculties need to be developed via environmental stimuli and interaction. Human neonates do not begin their development by conceptual abstraction and conceptual synthesis, although they will typically eventually develop this ability. Habermas’s use of Kohlberg’s theory of moral development in his own ethical theory is instructive of this.⁸⁶ The bi-conditional relationship, then, is sound. Conceptual synthesis and conceptual abstraction are impossible without logical judgment, and logical judgment is impossible without conceptual synthesis and conceptual abstraction.

Another objection to my respective conceptions of conceptual synthesis and conceptual abstraction might yet be in order. One might object that the way I have been describing these mental phenomena is scientifically inaccurate, either from a psychological or neuroscientific perspective. We do not, strictly speaking, form syllogisms in order to make judgments, infer cause-and-effect relationships via on-the-fly inductive and deductive logic, or identify objects by thinking of the law of identity and applying it to objects. My point here is not to begin a lengthy exegesis or analysis of the debate between Heideggerian versus Kantian or psychological

⁸⁵ Bandura, *Social Foundations*, 56.

⁸⁶ Jürgen Habermas, *Moral Consciousness and Communicative Action*, (Cambridge: The MIT Press, 1990): 170.

versus neurophysiological versus phenomenological interpretations of mental phenomena. I simply admit the following, very sensible, point: at a very basic level of cognition, what Kant terms sensibility, no sentence-form judgments operate in order to assess reality; as we have seen, it is a complex process of symbolization which achieves this. Even at these lowest levels of cognition, however, a sort of logic seems to be operative. As Andrew Brook's psychological analysis of Kantian epistemology suggests, the interaction between lower-level sensibility and higher-level understanding, and thus experience altogether, would be impossible without such an unconscious logical process.⁸⁷ Thus, in lieu of a comprehensive inquiry into the interaction and cognitive borders between sensibility and understanding, the terms that I use in my analysis with regards to syllogistic formation, identification, logical inference, and logical judgment more generally, form a convenient shorthand for more complex mental processes whose elucidation lies beyond the scope of this study. These terms, therefore, are not to be construed as exhaustive characterizations of mental phenomena which, in reality, cannot be explained exhaustively via the use of simple logical terms. The use of these terms, however, does have some legitimate purpose. I have used many of these terms with the intention of making explicit the links between important normative and descriptive concepts and the deep logical structure that is embedded within our cognitive-perceptual apparatus.

Rationality has, as I have been indicating, a vital experiential component. Bandura and Rottschaefer's account of agency lends much evidence to this. Logical judgment, and certainly any logical judgment that takes into consideration a multiplicity of complex variables, necessitates the participation of levels of agency above the base level. This participation of the various levels of agency in a reciprocally deterministic fashion in logical judgment necessitates freedom or voluntariness, or as a hardline determinist might insist, a sense of freedom or voluntariness. This is also bi-conditional, since freedom necessitates logical judgment, in at least the minimal sense required to make any sort of decision or deliberation. This bi-conditional relationship is one that Sen has also duly noted.⁸⁸ We can, thus, rewrite the above statement as:

$$\text{HasLogicalJudgment}(\Phi) \leftrightarrow (\text{HasConceptualSynthesis}(\Phi) \wedge \text{HasConceptualAbstraction}(\Phi) \wedge \text{HasFreedom}(\Phi))$$

We are now in a position to narrow our focus to two main conceptions of rationality relevant to the domain of ethics. The first I will call "logical rationality," and the second I will call "experiential rationality." Logical rationality is what people often refer to simply as "logic" or "rationality," often interchangeably and ambiguously; it is the rationality of rational-choice and social-choice theory. Ethical theorists often make the distinction between practical and theoretical rationality, both of which

⁸⁷ Andrew Brook, *Kant and the Mind*, (Cambridge: Cambridge University Press. 1994).

⁸⁸ Sen, *Rationality and Freedom*.

we can justifiably place under this category of “logical rationality.” This conception of rationality pertains to logical judgment itself and the specific activities that characterize it. These two subcategories of practical and theoretical rationality fit within this conception of rationality because they characterize specific modes of judgment: one concerning action and the other concerning the nature of reality. Within theoretical rationality, there are the further subcategories of deduction and induction. This distinction, however, between practical and theoretical judgment might turn out to be an artificial one, since action’s non-existence in reality cannot be meaningfully asserted. Nonetheless, such a distinction is still often practically significant. Practical rationality, as I and many others contend, also contains the subcategories of induction and deduction, although emotivists, intuitionists, pure scientific integrationists, many contractarians, and some others would disagree with the placement of one or both of them into this category. Many other characterizations of rationality have taken into account only this conception of rationality, and have run into major difficulties as a result. As we shall see, however, it is not logical rationality, by itself, that constitutes the essence of persons.

Although it is logical rationality that characterizes our judgments of the truth conditions or, at the very least, the consistencies of our judgments via induction and deduction, logical judgment must not be assessed *ex hypothesi*. As the essence of logical rationality, logical judgment is both necessary and sufficient for conceptual synthesis, conceptual abstraction, and freedom or voluntariness. The truth conditions which govern statements concerning the phenomenon of logical judgment, therefore, also govern those of conceptual synthesis, conceptual abstraction, and freedom or voluntariness. Experiential rationality, defined as those faculties that are necessary and sufficient for logical judgment, can also be seen to govern the totality of our experience as sapient beings. As might by now be obvious, the antecedent and consequent of the previous bi-conditional are the essences of logical rationality and experiential rationality, respectively. Thus, the bi-conditional can be re-expressed in the following way:

$$\text{HasLogicalRationality}(\Phi) \leftrightarrow \text{HasExperientialRationality}(\Phi)$$

What, then, is the *concept* of rationality? Isn’t this simply a representation of two different *conceptions*? The concept of rationality, can, in some sense, be understood by considering what the connective represents. It, in fact, represents a necessary relationship between these two conceptions. In conceiving of the antecedent, one necessarily conceives of the consequent, and in conceiving of the consequent, one necessarily conceives of the antecedent. It is this necessary relationship, therefore, that is the concept of rationality. Unlike our *conceptions* of rationality, which focus on one aspect of the relationship, when we focus on the relationship itself, we focus our attention toward the concept itself. To clarify this a bit, consider that *pros hen* equivocal uses of words which we use every day would be impossible without our focus, via the use of a word according to some specific interpretation, on some specific aspect of a common logical-linguistic relationship. Whenever we use a word according to some personal or institutional interpretation, it is always an

interpretation of *something*. That something is a necessary relationship between conceptions or, more simply, a concept.

Let us see, then, if rationality, as I have characterized it, lives up to the criteria I have previously given. The first condition to which I proposed that any definition of rationality must conform is as follows: “that it reflect common usage, that is every possible usage, of the term.” To try to poke holes in this definition, then, we must ask the question, “Is there any possible conception of rationality which could be defined in any terms other than the one given by the necessary relationship $\text{HasLogicalRationality}(\Phi) \leftrightarrow \text{HasExperientialRationality}(\Phi)$?” One might say that a decision is rational, that a patient is rational and can, thus, consent to treatment, that, conceivably, music is rational, and so on without straying from what I have proposed as the concept of rationality. The point is that all of these uses of “rational” have *pros hen* equivocal meaning in that they all implicitly refer to the aforementioned logical relationship. “Aha,” one might say, “Music is rational in a different sense than you are using the word, since music is always created and experienced from the heart.” While this might be true, is it not a necessary condition both of music composition and appreciation that the composer and listener have experiential rationality, that is, that they at least have the capacity for conceptual abstraction? Do not both have the capacity for symbolic communication and interpretation, as Hauser has characterized it? We must answer in the affirmative. Furthermore, does not Korsgaard’s Kantian conception, Gert’s hybrid conception, Aristotle’s teleological conception, Apel’s conception, game theorists’ and expected-utility theorists’ conceptions, and every possible interpretation of rationality that Macintyre has set forth conform to the concept of rationality as I have presented it? I submit that they, and every possible conception, do.

That leaves us with my second condition, which is as follows: “that [a definition of rationality] at least be consistent with the predictions and explanations provided by scientific models.” While it is a necessary condition of understanding, at least in part, what rationality is that we have a definition of rationality based upon our everyday experience and word usage, we must also realize that empirical observation has consequences for how we interpret these experiences and this word usage. Psychology and biology at least partially provide us with the tools which aid us in such a regard and which, in other words, help us to gain a fuller understand of that to which we are referring. As it should be evident by now, various perspectives in psychology, biological anthropology, and neurophysiology have been surveyed for this purpose. The fulfillment of this criterion, then, should be evident. Phenomenology, by contrast, which captures those parts of rationality which the sciences do not, will be surveyed more in depth as we progress.

2.3.6 Consequences for a Doctrine of Liberation

It is worth mentioning that, while I do not readily accept Critical Theory as an epistemological framework, there is a somewhat close parallel between what such

critical theorists as Horkheimer and Adorno have submitted as their interpretation of the distinction between *Verstand* and *Vernunft* and my distinction between logical rationality and experiential rationality. Logical rationality, like *Verstand*, stands in a certain role-relation to deconstruction (in a non-post-structuralist sense) and reduction. It is, as a critical theorist might say, the source of reification. Experiential rationality, like a Kantian interpretation of *Verstand*, is capable of employing both the preconditions and the products of logical rationalization and reduction in order to synthesize and construct a coherent, self-conscious experience. As such, experiential rationality also encompasses phenomenology. Experiential rationality, thus, truly does encompass the totality of our experience, even and especially those parts of reality which are largely constituted by relations of meaning or meaningfulness. While there are certain parallels between the critical-theoretical conception of rationality and my conception, *Verstand* and *Vernunft* are also significantly different from logical and experiential rationality, respectively. My focus in distinguishing between logical and experiential rationality is not particularly on the mathematically, logically, or paradigmatically reductive capacity of logical rationality; my aim is not to point out, as the critical theorists have, that a coldly logic-focused conception of rationality reduces experience and observation and thereby unduly compartmentalizes knowledge. Likewise, my aim is not to demonstrate, as the critical theorists also have, that such experience-focused conceptions as *Vernunft* are capable of expanding our capacity for a more holistic conception of knowledge. Rather, my aim is to demonstrate, through parallel with the critical-theoretic conceptions of rationality, that logical and experiential rationality, as mutually nourishing and reinforcing entities, yield two key ideas embedded within them that theorists such as Habermas have also highlighted: ideological self-criticism and individual liberation via the holistic epistemological and meta-physical appraisal and reappraisal of oneself. The idea of freedom inherent in experiential rationality, therefore, is free from various cultural conceptions; it is, as it were, the pure concept of freedom. Some philosopher-anthropologists might scoff at such an idea, but, nevertheless, I believe this is a true claim. Thus, the idea of liberalism, tainted as it has been by what Herbert Marcuse has rightly characterized as a “one-dimensional” conception of freedom, might be inadequate to help us in promoting the highest form of liberty, if indeed liberty is worth seeking at all. Apel himself points to such a problem inherent in the liberal conception of liberty via what he terms “the *conventionalist-liberalistic fallacy* of confusing freedom of moral conscience with private arbitrariness of decision.” As he states:

The intrinsic fallacy of that *conventionalism* which obscures the basic deficiency of the (Western) complementarity-system lies, in my opinion, in the fact that freedom of conscience as moral autonomy is confused with private arbitrariness of decision. Thus it is overlooked that when the so-called free decisions of conscience are no longer based on intersubjectively valid (binding) norms, they must automatically fall prey to external causal determinations, be they dispositions of one’s instinctual nature, or social manipulations.⁸⁹

⁸⁹ Apel, “Types of Rationality Today,” 142.

Thus, one of the pictures that will slowly emerge throughout the rest of this work is that the liberal conception is not enough; it is indeed a very shallow and naïve conception of liberty. What we will need, instead, is a libertarianism, or even more precisely, a doctrine of liberation.

2.4 Concluding Remarks

Where, then, do we go from here? Certainly, rationality is centrally relevant to ethics, but how can we draw anything action- or character-directing from it alone? In the next three sections, I will present a partial answer to this question. In Sect. 3.1, I will vindicate naturalism against Hare's objections to it, and I will also give an account of Hare's universal prescriptivism and preference-utilitarianism. I will argue that his theory has less far-reaching application than that which he has purported. In Sect. 3.2, I will use the part of Hare's theory which I believe to have application to ethics, along with some concepts discovered by other theorists, in developing a method of ethical justification. These two sections demonstrate what role logical judgment, alone, has in ethics. In Sect. 3.3, I will expand on the method of justification, showing, through a brief analysis of Gewirth's ethical theory, how experiential rationality will eventually come into the picture. Gewirth's theory, centered around the phenomenon of agency itself, and especially the judgments one necessarily makes about one's own agency, demonstrates that logical rationality is far-reaching in its action-directing scope. In Sect. 3.4, I will demonstrate that Habermas's discourse theory has much the same consequences and, indeed, that it even extends the applicability of Gewirthian theory. Although extremely successful, Gewirth and Habermas's theories, perhaps because they do not fully take into account scientific evidence or phenomenology (the dearth of phenomenological analysis applies mostly to Gewirth), do not take into account all of the conceptual and logical consequences of the concept of rationality, namely important components of experiential rationality. In order to elucidate these consequences fully, and thus demonstrate how deontological ethics finds grounding only in virtue ethics via a Footian-Carnapian "grammar of goodness," it will be necessary to read the last two chapters on deontic concepts and goodness, respectively. At that point, we shall leave the world of mere logical-judgment-based ethics and enter the world of that based upon the holistic synthesis of both logical and experiential rationality. This, as we shall see, will have significant consequences for value theory.

Chapter 3

Rationality and Dialectical Necessity

As the name of this chapter implies, I apply my previous characterizations of rationality, rationalism, and naturalism to the prominent dialectical ethical theories of Hare, Gewirth, and Habermas. All three theories share important commonalities with regards to pragmatic consistency and practical justification. These commonalities constitute a major strength in their theories and thus actually offer a powerful collective dialectical model, but their accounts fail to address and incorporate many important features of rationality as I have characterized it and, as such, fail to recognize the meta-theoretical basis for aretaic and teleological conceptions of action, behavior, and general states of affairs. Not only does the dialectical approach inherently de-emphasize the theoretical consequences of experiential rationality and thereby fail to have application to what might be considered *general* questions of value (such as the questions cited at the beginning of this book, as well as even such questions as “What is good?” and “What is better?”), but its dialectical conception of logical judgment is also limited by its depiction of ethical judgment as purely dialectical. I show, therefore, that there are major semantic and conceptual gaps in this theoretical framework with regards to definitions of “good,” “better,” “ought,” and “must” and that, without accompanying analyses and characterizations of these concepts, certain major components of their theories are exposed as vapid and lacking in directional content. It can be shown, I argue, that, because such theoretical frameworks overlook such analyses, they miss a vital foundational component of their own theories, a component that is inherently aretaic in nature.

3.1 Prescription, Preference, and Dialectical Contingence

In this section, I analyze and critique the dialectical theory of R. M. Hare. I argue, as Gewirth does, that Hare’s dialectically contingent approach is highly problematic. I discover, however, important elements of pragmatic consistency that contribute to a negative justificatory approach. I also take this space to undermine some of Hare’s

main and most influential arguments, namely those on desiderative internalism of value and anti-descriptivism. As the premises of these arguments have heavily influenced all subsequent argumentation in favor of these positions, their refutation will prove integral to the firm establishment of the theoretical grounds I will propose. I will ultimately conclude that the one vital component of his theory to an overall theory of value is its relevance to negative pragmatic justification. As such, I go on to use his and some other non-dialectical approaches in order to demonstrate that some ethical statements are *a priori* false. In this section, therefore, my goal is two-fold: (1) to vindicate naturalism against non-naturalist and anti-descriptivist claims and demonstrate, deductively, that some form of naturalism must be sound and (2) to demonstrate the applicability and contribution of R.M. Hare's dialectical account to a negative theory of justification. The achievement of these goals, and in particular the establishment of a negative justificatory approach, will allow me, without first presupposing the soundness of a particular positive account of justification and thereby leaving my argument open to the criticism that the conclusions follow only on the acceptance of my own account, to clear such positions as ethical egoism, relativism, subjectivism, nihilism, and some forms of utilitarianism from consideration as viable theories. This, I believe, is the most reliable method of justifying the preliminary rejection of certain theoretical frameworks and approaches. This, in turn, will give us a firm foundation for the formation of a far-reaching positive theory of ethical justification.

Hare has had an extraordinary influence on the fields of ethics and meta-ethics, especially for the rationalist paradigm. Along with Philippa Foot, to whom he is opposed, Hare successfully challenged the claims of such emotivists as Charles Stevenson and such subjectivists as Rudolf Carnap by demonstrating that ethical judgments, like language in general, follow logical rules. Having modeled much of the structure of his theory on the Kantian categorical imperative, he has claimed that, if we are careful about how we apply the principle of universalization, we do not, as Kant insists, arrive at a rigid deontology that takes no account of the peculiarities and contingencies of specific situations. He claims, rather, that, given the correct use of the principle of universalization, we arrive at what he terms "preference-utilitarianism." If preference-utilitarianism does follow from the premises of his theory, then we might confidently state, as Hare does, that it avoids the problems inherent in other forms of utilitarianism. It is not at all clear, however, that such a utilitarianism does follow or even that the premises of his theory are sound.

Hare's account, we must admit, is a thorough and thoughtful one, but, as we shall see, it remains mired in some serious issues. Although flawed in some serious ways, the theory is not fatally flawed; it simply does not yield the conclusions Hare claims it does. Since his has been such a momentous and influential undertaking, and since it has substantial consequences for rationality and a dialectical method of justification, it would be an major oversight not to analyze and consider his account. As I will demonstrate, there are two main issues in Hare's theory that need to be addressed. The first is his vehement opposition to any theory that is remotely naturalistic or descriptivist. I shall examine his grounds for such opposition and determine whether they are sound. As I will demonstrate, not only are his grounds for such opposition

unsound, but so are the grounds of most, if not all, arguments against naturalism and descriptivism. The second issue to be addressed is the manner in which Hare attempts to justify his ethical theory. As I will demonstrate, his use of the principle of universalization, on any interpretation, fails to yield substantial, justifiable ethical norms. It is important to note, however, that, if Hare reformulated his theory to limit the scope of criteria of relevant similarities to those prescriptions which make explicit reference to commonly held preferences, instead simply of momentarily or occasionally held preferences, his theory would have important theoretical consequences, namely that, through an implicit link to Universal Moral Grammar Theory, certain grounds for ethical judgment and justification could be decisively shown to be unsound via the dialectical instantiation of certain pragmatic inconsistencies.

3.1.1 A Refutation of Harean Anti-descriptivism

The first issue to take up, then, is Hare's opposition to naturalism and descriptivism in general. It is important to note, at the outset, that the bulk of his objections to naturalism have evidently, and in some instances obviously, been directed at Foot, whose positions on naturalism have changed and evolved during the course of her career. The best representation of the culmination of her naturalist views can be found in her most recent work, *Natural Goodness*.¹ Thus, she has addressed many of his criticisms. I believe, however, that she has not satisfactorily answered all of them. It will be important, then, to press more seriously against Hare's claims.

Next to G. E. Moore, Hare is possibly one of the greatest proponents of the Open-Question Argument of the twentieth century. In his *Principia Ethica*, Moore, in his argument for the non-natural status of goodness, set out what he took to be a fatal flaw in theorists' attempts to define "good" in terms of some natural property. This position is best captured in the following passage:

What we want to know is simply what is good. We may indeed agree that what most people do think good, is actually so; we shall at all events be glad to know their opinions: but when we say that their opinions about what is good, we do mean what we say; we do not care whether they call that thing horse or table or chair, gut or bon or ἀγαθός; we want to know what it is that they so call. When they say Pleasure is good, we cannot believe that they merely mean Pleasure is pleasure and nothing more than that.²

In other words, on Moore's account, we seem to be lacking some element of meaning when we attempt to define "good" in terms of a natural phenomenon such as pleasure, since, when we make such statements as "X is good because it is pleasant," we can ask the question, "But why is pleasure good?" On Moore's account, any attempts to prove that pleasure is good, say, by retorting, "Because pleasure makes us feel good, contributes to social cohesion, etc.," are doomed to failure because we will

¹ Philippa Foot, *Natural Goodness*, (Oxford: Oxford University Press. 2001).

² G. E. Moore, *Principia Ethica*, (Cambridge: Cambridge University Press. 1903): 11.

always be lacking the central meaning of “good.” In other words, he is claiming that we are not defining “good” at all; instead, we will forever simply be describing more good things and will never explain what makes them good and why.

Hare takes up this line of reasoning as well throughout his career, and with vigor. Hare’s account of this supposed problem is very similar. As he explains:

If . . . we wish to say that the members of the Royal Academy have good taste in pictures, we have, according to some definition, to say something which means the same as saying that they have this feeling of admiration for pictures which have a tendency to arouse in them this feeling. This is not what we wanted to say. We wanted to say that they admired good pictures; we have succeeded only in saying that they admired pictures which they admired. Thus if we accept the definition we debar ourselves from saying something that we do sometimes want to say.³

Hare goes on to take the example of the judgment “S is a good strawberry.” We might be tempted, he claims, to state that what we are saying is something along the lines of “S is a good strawberry because it is sweet, juicy, firm, red, and large.” This, he asserts, is problematic because this does not capture the full meaning of what we say in normal speech. He explains, “We sometimes want to say that a strawberry is a good strawberry because it is sweet, &c. This. . . does not mean the same thing as saying that a strawberry is a sweet, &c., strawberry because it is sweet, &c.”⁴ This, he explains, is clearly not what we wanted to say. This course of argumentation, however, is deceptive.

There have been biting counterarguments to the Open-Question Argument. Take, for example, David Brink’s argument on the presupposition of predicative use in the statement “X is good.”⁵ Alternatively, one might consider William Frankena’s objection on the grounds that the argument presupposes analyticity.⁶ Although Frankena’s objection to the argument has received criticism on internalist-reasoning grounds, this chapter will adequately deal with the problems inherent in such a model of reasoning. One might also consider Arthur Prior’s objections,⁷ Rawls’s objections,⁸ or, more recently, those of Michael Smith⁹ and even Rottschaefer.¹⁰ Hare, however, in none of his main works, directly takes up a single one of these many objections although, to be fair, Smith and Rottschaefer’s objections appeared after he wrote his final work.

Still, riding somewhat on the coattails of other theorists, I have a few objections of my own. First, let us consider the main claims of the argument, namely that some sort of meaning is lost that we can never paraphrase or, otherwise, capture in

³ R. M. Hare, *The Language of Morals*, (Oxford: Oxford University Press. 1952): 84.

⁴ *Ibid.*, 85.

⁵ David Brink, *Moral Realism and the Foundations of Ethics*, (Cambridge: Cambridge University Press. 1989): 52.

⁶ William Frankena, “The Naturalistic Fallacy,” *Mind*. 48 (1939): 464–477.

⁷ Arthur N. Prior, *Logic and the Basis of Ethics*, (Oxford: Oxford University Press. 1949).

⁸ Rawls, *A Theory of Justice*.

⁹ Smith, *The Moral Problem; Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics*, (Cambridge: Cambridge University Press. 2004).

¹⁰ Rottschaefer, *The Biology and Psychology of Moral Agency*.

ordinary language. If some deeper, hidden meaning is lacking in our attempts to define “good,” then one would suppose that non-descriptivists could tell us what, in some capacity, this missing meaning is. The typical non-descriptivist response to this might be, “But that is precisely my point: the missing meaning cannot be expressed through language.” If this is the case, however, then it must at least be some emotive meaning that is missing. This thesis, however, also fails; both Hare and Foot, as well as others, have amply demonstrated the flaws of the emotivist position.¹¹ This leads us back to the original problem: if there is some missing meaning, non-naturalists and non-descriptivists should be able to tell us what it is, or at least be able to characterize it in some non-circular way. The non-descriptivist, however, denies this, and herein lie the great problems with non-descriptivism. On the first account, the non-descriptivist tells us that there is some missing meaning in our attempts to paraphrase or define “good,” yet on the second account, she tells us that she cannot characterize this missing meaning in any non-circular way. She cannot have it both ways. If the non-descriptivist insists on claiming that there is some missing meaning, she cannot simultaneously claim that she does not know, in any sense, what this missing meaning is. How can one make a claim that something is the case and also make the claim that she doesn’t know what that something is, especially when we are speaking of meaning, which, by definition, means *something*? Hare tries to get off the hook by saying that this missing meaning is “evaluative” or that statements of goodness express “standards” and “commendation.” This is mere obfuscation. We might, likewise, pose the non-descriptivist with her same open question. “What,” we might ask, “is evaluation?” The concept of evaluation, of course, entails judging something to be better or worse than something else; it entails, in fact, a judgment of goodness. In a struggle not to refer to goodness, the non-descriptivist might offer us the answer, “Evaluation is the setting up of a standard or the act of commending.” This is yet further obfuscation, for we can, yet again, pose the open question, “What is a standard?” or “What is commendation?” The fact is that standards are the result of evaluations, which I previously suggested are definable only in terms of goodness. Commendation, likewise, is to think or judge something to be good. Even to state that commendation is to judge positively is to say that one has some sort of attitude, even some feeling, toward some object such that one would ascribe value to it, and even if we speak fallaciously of value as “warm feelings” and the like, this term “warm” is merely a cover for stating that these feelings themselves are valued by the agent, that is, thought to be valuable. In other words, as Moore himself states, the agent

¹¹ Brand Blanshard, *Reason and Goodness*. (Atlantic Highlands: Humanities Press. 1961); Philippa Foot, “Moral Beliefs,” *Proceedings of the Aristotelian Society* (1958–59), 59:83–104. Reprinted in Philippa Foot, ed., *Theories of Ethics*, pp. 83–100; Hare, *The Language of Morals; Freedom and Reason*, (New York: Oxford University Press. 1965); *Moral Thinking: Its Levels, Method and Point*, (Oxford: Oxford University Press. 1981); *Sorting Out Ethics*, (Oxford: Oxford University Press. 1997); Macintyre, *After Virtue*; Thomas Nagel, *The Last Word*, (Oxford: Oxford University Press. 1997); Sen, *The Idea of Justice*; Henry Babcock Veatch, *Rational Man: A Modern Interpretation of Aristotelian Ethics*. (Indianapolis: Liberty Fund. 2003).

would not merely speak of this warmth as pleasant or simply as warmth; he would provide *reasons* or a *rationale* for why he should continue to feel such warmth; he would, in other words, describe the warmth as good in addition to describing the fact that the feelings are indeed warm. From this perspective, we can see that Hare's reasoning is self-defeating, for let us take one of Hare's principal claims about statements of goodness: "To judge something good is to commend it." If, as I suggest, we can only define commendation in terms of goodness, then we end up with the statement, "To judge something good is to judge it to be good," which mirrors a central criticism that non-descriptivists and non-naturalists have hurled at naturalists, namely that (1) naturalists cannot define "good" in any non-circular way and that (2) "good" is literally meaningless while such terms as "evaluation," "standard," "commendation," and "pro-attitude" are not. The further implied claim here is often that "good" then has emotive meaning. As we have seen, however, non-descriptivists have failed to characterize this emotive meaning without referring to evaluations, standards, commendations, and pro-attitudes as emotive, which clearly begs the question. The same type of open question can therefore be applied. "But why does 'good' have a commendatory function?," we might ask. "Because 'good' is evaluative," the non-descriptivist will respond. "But why is 'good' evaluative?," we may further ask. "Because 'good' is an emotive term." "But why is 'good' and emotive term?," we will surely press. "Because evaluation is emotive," the non-descriptivist might respond. "But why is evaluation emotive?," we ask. "Because evaluation is non-literal, and it has missing meaning," he might respond. "But what is this missing meaning?," we will ask. "The missing meaning is emotive," the non-descriptivist might tell us. As we can see, due to the absurdity inherent in the non-descriptivist responses and the circularity of his arguments, such a dialogue can go on forever. Thus, we find that the Open-Question Argument is founded on nothing but linguistic confusion. Based on our observations, both social-scientific and interpersonal, we can also see that, while persons might often act on emotion, norms do not acquire meaning through emotion alone. As Bandura's theory, as well as a myriad of psychological and sociological research, demonstrates, individuals utilize reflective norms or systematic rationales, often both from peer pressure or individual reflection, in order to carry out their actions.¹²

¹² Catherine E. Amiot, et. al., "Can Intergroup Behaviors Be Emitted Out of Self-Determined Reasons? Testing the Role of Group Norms and Behavioral Congruence in the Internalization of Discrimination and Parity Behaviors," *Pers Soc Psychol Bull* January 1, 2012 38: 63–76; Bandura, *Social Foundations; Self-Efficacy in Changing Societies; Self-Efficacy: The Exercise of Control*; Seln Kesebir, "The Superorganism Account of Human Sociality: How and When Human Groups Are Like Beehives," *Pers Soc Psychol Rev* August 1, 2012 16: 233–261; N. Kugihara, "Effects of aggressive behaviour and group size on collective escape in an emergency: a test between a social identity model and deindividuation theory," *Br J Soc Psychol*. 2001 Dec;40(Pt 4):575–98; Kim-Pong Tam, et. al., "Intersubjective Model of Value Transmission: Parents Using Perceived Norms as Reference When Socializing Children," *Pers Soc Psychol Bull* August 1, 2012 38: 1041–1052; Walter Mischel and Yuichi Shoda, "A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure," *Psychological Review* 1995, Vol. 102, No. 2, 246–268.

To state that these norms are reflective means that, although they might not always manifest in conscious thought, they will either be defended or rejected by the agent upon the basis of *some* sort of rationale. As Bandura affirms: “Social cognition theory underscores the self-arousal power of cognition. Indeed, physiological arousal itself is often generated cognitively by arousing trains of thought.”¹³ Thus, he emphasizes, “To the extent that people can regulate what they think, they can influence how they feel and behave.”¹⁴ Thus, while it might even be true that much norm formation occurs as a result of emotive processes, the agent has the ability, unless certain mental illnesses are present, to choose whether or not to accept the norm or rationale. In other words, even if emotion plays a role in the formation of norms (which it likely does to some extent), then it does not follow that the meaning of the norm for the agent is purely emotive. Instead, the meaning of the norm is deliberative; the formation of the norm as legitimate from the perspective of the agent can be promoted or curtailed upon reflection. This is primarily salient in the fact that emotion does not determine cognition. As Bandura explains:

One can distinguish three principal ways in which efficacy beliefs affect the nature and intensity of emotional experiences: through the exercise of personal control over *thought*, *action*, and *affect*. The thought-oriented mode in the regulation of affective states takes two forms. Efficacy beliefs create attentional biases and influence whether life events are construed, cognitively represented, and retrieved in ways that are benign or emotionally perturbing. The second form of influence centers on perceived cognitive abilities to control perturbing trains of thought when they intrude on the flow of consciousness. In the action-oriented mode of influence, efficacy beliefs regulate emotional states by supporting effective courses of action to transform the environment in ways that alter its emotive potential. The affect-oriented mode of influence involves perceived efficacy to ameliorate aversive emotional states once they are aroused. These alternative paths of affect regulation are amply documented in the exercise of control over anxiety arousal, depressive mood, and biological stress reactions.¹⁵

It cannot, therefore, be the case that norms and reasons, being cognitive entities, arise from and are determined by emotion. Agents adopt *reasons* for their actions. It is these reflective, deliberative reasons, and not mere emotion, that plays a significant role in identity formation, self-efficacy, and ultimately the way in which individuals interact within their social milieux. This process of rationalization and reflective evaluation, in turn, forms a vital part of how agents necessarily conceptualize and utilize the concept of goodness, and this conceptualization thus necessarily shapes the meaning of “good.” While I have argued against the non-descriptivist thesis deductively, therefore, we can also see that it is likewise inductively implausible that “good” is a merely emotive term.

A behaviorist might suggest that that to which we refer when we make statements of goodness can be explained using the concepts of attractors and repulsors: when we think something good, we are attracted to it, and when we think something

¹³ Bandura, *Self-Efficacy: The Exercise of Control*, 139.

¹⁴ *Ibid.*, 145.

¹⁵ *Ibid.*, 137.

bad, we are repulsed by it. There is only one problem with this: agents are not moths. We do not simply see a “light” and move toward it non-cognitively. Even if the majority of our actions were compulsive or unintentional, it would take but one instance of an intentional judgment of goodness to demonstrate the falsehood of this thesis. Furthermore, it is not the case that, when we judge things to be good or bad, we are attracted or repulsed by them in the sense that the behaviorist defines these terms. I might think broccoli is good for me but not be motivated to eat it; this would take a metalevel motivational impetus. It would, in other words, take a *reason*.

The reader might, however, remain skeptical. Have I not, indeed, proved the non-descriptivist right? Have I not precisely demonstrated that a definition of “good” is impossible in terms of anything else? It might be tempting to say that I have, but I have not demonstrated this; I have simply demonstrated that I have not yet proved “good” to be definable in terms of natural properties and, furthermore, that the non-descriptivist has not proved it to be indefinable in terms of natural properties. The burden of proof, however, is currently on the non-descriptivist to demonstrate such indefinability, since I have demonstrated that his arguments against naturalism and for non-descriptivism are the same sorts of grounds that can be used to invalidate non-descriptivism. Since the non-descriptivist thesis is subject to such counterexample, the non-descriptivist argument is invalid. We can, in fact, take this one step further. If these are the grounds for any non-descriptivist argument, and these grounds are invalid, does it not follow from this that, if the only alternative is descriptivism, that some form of descriptivism must be sound? In order to press this a bit, let us pose an important question to the non-descriptivist: “If goodness is not a property, then what is it?” If it is a mere logical operator, deontic or otherwise, non-descriptivists will have to demonstrate this. If it is not a logical operator, then the word refers to nothing, which is certainly extremely odd. How can there exist a word that refers to nothing, unless it is meaningless? And if “good” is meaningless, then we seem to be stuck right back in some form of emotivism or intuitionism, which also begs the question. Allow me to explain why this is the case.

Magnell has explored the predicative-attributive distinction with regards to adjectives, in general, that signify properties.¹⁶ Such adjectives as “wide,” “large,” “short,” and “soft” are attributive, since, contained in their meaning is an element of comparison. Something is wide, we say, only if it has a certain width in relation to some other object that has some certain width. The adjective, although containing an implicit reference to an intrinsic property, namely width, refers to a property that is instantiated extrinsically by comparison to other objects that have the same intrinsic property. The adjective “good” adheres to the same linguistic pattern. If this is the case, then we can apply the Open-Question Argument to every

¹⁶ Thomas Magnell, 1993a, “Evaluations as Assessments, Part I: Properties and Their Signifiers,” *Journal of Value Inquiry*, 27: 1–11; 1993b, “Evaluations as Assessments, Part II: Distinguishing Assertions and Instancing Good of a Kind,” *Journal of Value Inquiry*, 27: 151–63.

attributive adjective in our language. “What makes X wide?,” we might ask. Finding an answer to this seems equally difficult. We might be tempted to respond “X is wide because it has a greater width than Y,” but, of course, we are begging the question, “What is width?” It seems that, in order to define the extrinsic property, we must first define the intrinsic property. This, however, can be done; width is defined in terms of a set of spatial relationships; it is a property of space-time. Would it not be bizarre, then, that every other adjective in our language can be adequately described, characterized, or defined, but that “good” is mysteriously exempt from this? It would certainly be bizarre that, when using adjectives, we make reference to some existing entity, whether or not it is of an emotive nature, but that, when we use the word “good,” we refer to absolutely nothing. Despite the claims of Error Theory and J. L. Mackie’s Argument from Queerness,¹⁷ this is what would seem to be the real Argument from Queerness, namely that we refer to something when we use adjectives but refer to nothing when it is the adjective “good.” This is akin to the sort of view which Mark Schroeder seems to suggest in his elucidation of the Frege-Geach Problem.¹⁸ We must not, however, dismiss positions based on their queerness; we must, rather, question the *intelligibility* of any claim that it is possible to refer to nothing in our use of adjectives. While not addressing this particular problem directly, Smith has argued, in opposition to McDowell, that Mackie’s Error Theory retains considerable credibility due to certain metaphysical questions it raises concerning the phenomenology of value.¹⁹ Smith, however, seems to be pointing to a deeper metaphysical debate, not strictly a meta-ethical one, concerning the nominalist position with regards to secondary qualities. In other words, if Smith’s account of Mackie’s thesis is correct, then all secondary qualities that could conceivably be subject to dispositional analysis, including color, heat, odor, and a host of others, and not simply value, would not be “there to be experienced” in reality. The supposed tension he points out between dispositional analysis and the “explanatory test” with regards to secondary qualities and value, therefore, essentially falls back on classic metaphysical debates over nominalism and materialism. Although I do take the position that qualia such as color, heat, and odor cannot be reduced to mere descriptions of physical phenomena, although they have definite correspondence to physical reality, this is irrelevant since, as I in any case previously stated, Mackie’s argument is essentially a metaphysical one which, if correct, would apply to all secondary qualities anyway. As such, I believe Smith’s counterargument to be moot. In consideration of the observations I have offered to this point, therefore, there

¹⁷ J. L. Mackie, *Ethics: Inventing Right and Wrong*, (London: Pelican Books. 1977): 38.

¹⁸ Mark Schroeder, “What is the Frege-Geach Problem?” *Philosophy Compass*. 3/4, (2008): 703–720.

¹⁹ John McDowell, “Values and Secondary Qualities,” *Mind, Value, & Reality*. Ed. John McDowell. (Cambridge: Harvard University Press. 1998): 131–151; Michael Smith, “Objectivity and Moral Realism: On the Significance of the Phenomenology of Moral Experience,” *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics*. Ed. Michael Smith. (Cambridge: Cambridge University Press. 2004): 234–258.

does not seem to be a naturalistic fallacy at all. If anything, the problem is one concerning a non-descriptivist fallacy.

Hare's anti-descriptivist arguments, however, do not cite solely the Open-Question Argument. Let us, therefore, address some of his other objections to descriptivism and naturalism. One of his criticisms concerns syllogistic structure. Hare states that naturalism fails because it often inserts value words into premises, thereby not truly satisfying the naturalistic condition that one begin with a naturalistic premise. "Nearly all so-called 'naturalistic definitions'," he asserts, "will break down. . . for to be genuinely naturalistic a definition must contain no expression for whose applicability there is not a definite criterion which does not involve the making of a value-judgement."²⁰ He purports, then, a dilemma: either practical syllogisms must be comprised of exclusively naturalistic premises and, thus, lack a prescriptive conclusion, or they must contain an imperative premise and yield a prescriptive conclusion but cease to be naturalistic.

Let us begin with the first part of his purported dilemma. If the syllogism is a naturalistic one, all of the premises must be naturalistic as well. Let us take, as Hare does, a hypothetical premise with some predicate C. "If the definition satisfies this test," he states, "we have next to ask whether its advocate ever wishes to commend anything for being C. If he says that he does, we have only to point out to him that his definition makes this impossible. . ."²¹ because the syllogism in question must be a naturalistic one. Hare's reasoning, however, is circular. He sets out to demonstrate that "commendation" and "value-words" cannot be part of the premises of any naturalistic syllogisms, for this, he asserts, would deprive the premises of any naturalistic status. In order to prove, however, that "commendation" and "value-words," which are in any case judgments of goodness, cannot comprise any part of naturalistic syllogisms, he would have to prove that such expressions are non-descriptivist in the first place. This sort of proof, however, on Hare's account, can only be obtained by presupposing that judgments of goodness are not naturalistic and by applying his rule that naturalistic syllogisms must not contain "value-words." Thus, in order to prove that statements of goodness are not naturalistic and, thus, that naturalistic syllogisms must not contain statements of goodness, he sets up the rule that "commendation" and "value-words," once again statements of goodness, are not naturalistic. Thus, without noticing it, Hare is stuck in a vicious circularity.

The second fork of his purported dilemma is equally insidious. His assertion, here, is that, if a syllogism contains an imperative premise, it is not a naturalistic premise and, therefore, is not part of a naturalistic syllogism. Again, we see a vicious circularity. He assumes that imperative statements are not naturalistic statements, covertly indicative, entailing an indicative, or otherwise, and he offers very little proof to the contrary. Thus, in order to prove that syllogisms containing imperatives are not naturalistic, he would have to prove that imperatives are not, on

²⁰ Hare, *The Language of Morals*, 93.

²¹ *Ibid.*

the same grounds, at least implicitly, naturalistic, and he has not done this. The little he offers to explain his thesis revolves around the following claim:

A statement, however loosely it is bound to the facts, cannot answer a question of the form ‘What shall I do?’; only a command can do this. Therefore, if we insist that moral judgements are nothing but loose statements of fact, we preclude them from fulfilling their main function; for their main function is to regulate conduct, and they can do this only if they are interpreted in such a way as to have imperative or prescriptive force.²²

Curiously, in *The Language of Morals*, he undertakes only a minimal analysis of imperative statements with the purpose of demonstrating that they cannot be derived from indicative statements. For instance, what, exactly, *is* a command? As with “commendation” and the like, Hare never directly addresses this. In Chap. 4, I will consider this point more in depth and offer an analysis which precisely demonstrates that imperatives can, indeed, be derived from factual statements. In addition to the above demonstrations of circularity, both forks of this purported dilemma hinge on the soundness of the Open-Question Argument, which, as I have previously argued, is invalid. The problems with his non-descriptivist thesis persist in his other works as well.

In *Freedom and Reason*, Hare sets out a thesis on descriptive meaning. He begins with an account of literal meaning in general. Literal meaning, he states, “is or involves the use of an expression in accordance with certain rules; the *kind* of meaning is determined by the *kind* of rules.” By rules, he explains, he means “consistency of practice in the use of an expression which is the condition of its intelligibility.”²³ He, thus, derives from his account of meaning that “In general, a person is misusing a descriptive term if in using it he breaks the descriptive rule attaching the term to a certain kind of objects; and he does this if he says that an object is of one kind, meaning, or intending to convey, that it is of another kind.” Therefore, he asserts, “A descriptive term may thus be defined as one, to misuse which is to do this.”²⁴ Already, however, the argument becomes mired in confusions. The way in which Hare uses, and goes on to use, the category “descriptivism” throughout *Freedom and Reason* does not, as he would like to suggest, exclude naturalism as a viable position. Again, his thesis rests on a faulty assumption: that “descriptive” and “evaluative” signify two exclusive categories. It is, perhaps, because he fails to rigorously analyze a vital component of meaning, namely the concept of reference, that he assumes these two categories follow two relevantly different sets of rules. Magnell, among others, has cast doubt on this. He has demonstrated that, although evaluations are often assessments, that is they pick out extrinsic properties via the use of attributive adjectives, it is inaccurate to state that evaluations, in general, are assessments.²⁵ Moral philosophers, he states, commonly misconstrue characteristics belonging to assessments as characteristics belonging to all evaluations. Magnell,

²² Ibid., 46.

²³ Hare, *Freedom and Reason*, 7.

²⁴ Ibid., 8.

²⁵ Magnell, “Evaluations as Assessments, Part I”; “Evaluations as Assessments, Part II.”

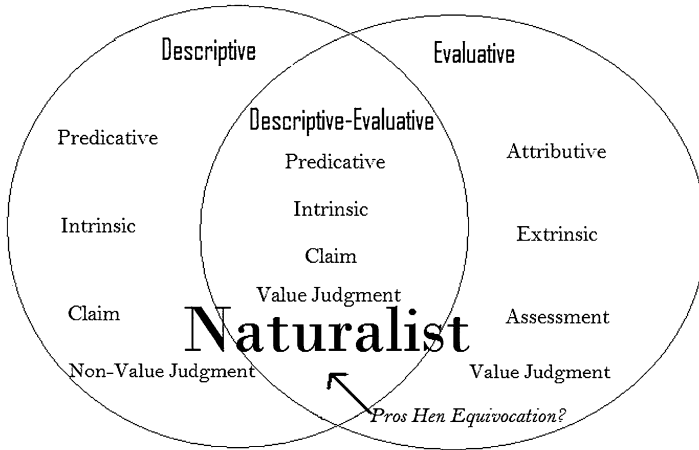


Fig. 3.1 Proposed classification scheme for descriptive and evaluative judgments

however, shows this to be false through repeated examples in which this is clearly not the case. We must, then, question Hare's characterization of the descriptive-evaluative distinction. If Magnell's account is correct, then there is the glaring possibility that there is an evaluative manner in which "good" can be used that is also predicative. This, then, would mean that the categories "descriptive" and "evaluative" are not exclusive categories with accompanying exclusive linguistic rules. On this account, "descriptive" and "evaluative" would be subsets of a larger category and that such a designation as "descriptive-evaluative" would be permissible. As represented in Fig. 3.1, naturalism could conceivably fall within either the "descriptive" or the "descriptive-evaluative" category. In addition, we cannot fail to consider that evaluative judgments that are assessments might implicitly denote descriptive-evaluative judgments via *pros hen* equivocation. Thus, all three categories might accommodate naturalist models. By contrast, it is unclear where a non-descriptivist model would fit into this picture, as all three categories accommodate explicit or implicit descriptions of natural properties. The essential difference between the "descriptive" category and the "descriptive-evaluative" category would seem to be that the latter category accommodates value judgments or, to be more specific, judgments of goodness and that the former does not. Note, however, that the predicativity and intrinsicity denoted by judgments in this category fulfill the criteria for factuality. Thus, Hare and other non-descriptivists have made two grave errors. First, they have assumed that "descriptive" and "evaluative" are two exclusive and distinct categories. This has led them to their second error, namely that facts and values are distinct. As we can see from Fig. 3.1, however, the descriptive-evaluative category is where fact meets value. In addition Hare and other non-descriptivists have also assumed the fact-value distinction due to an invalid inference that, since many value judgments belong to the evaluative category and because many of them are assessments, they must all belong to this category, a point which Magnell has refuted.

The basic error, then, is in assuming that evaluative statements can never be descriptive or vice versa.

In *Moral Thinking*, Hare offers yet more objections to descriptivism and naturalism. This time, they are on grounds of a supposed incongruity between language and culture. He explains:

It is not possible to treat the following as the single canon of moral reasoning: find out what sorts of things the moral predicates can properly be applied to according to the conventions of our language, and apply them only to those things. For this would be to treat the conventions as merely linguistic conventions, and as binding, therefore, on anybody who wishes to speak the language correctly. The mistake of confusing moral with linguistic conventions is analogous to the mistake of confusing moral with linguistic intuitions. Both have the effect of tying our moral reasoning to the received opinions of our society. It is an important feature of moral language, neglected by naturalists, that we can go on using the moral words with their same meanings to express moral opinions at variance with the received ones, as moral reformers do. This would be impossible if the moral words were tied by virtue of their very meanings to fixed properties of actions, etc.²⁶

This is a gross oversimplification of naturalistic claims. Essence, as Oderberg emphasizes, is not language-dependent.²⁷ Statements containing words with specific evaluative content are not, therefore, on the naturalist account, all true. Using, as Hare does, the example of racial slurs, it does not follow, from “I judge that he is an X,” where X is some racial slur, that “He is an X” is true. This is because, whatever evaluative meaning is contained in “X,” there had to have been, in the first place, an initial evaluation or set of evaluations to establish that word’s specific evaluative content. Since racial slurs are all the result of seriously flawed evaluations, we are at liberty to re-evaluate them. Drawing, yet again, upon the distinctions set out in Fig. 3.1, we might think of the difference between description and descriptive-evaluation as organized into two *layers* of facts. While the fact, for instance, that someone is of a certain skin color or geographic origin remains unchanged by such re-evaluations, one might think that what follows from this is that the supposed fact that this person is an X has also changed. This, however, is a mistake. The descriptive-evaluative fact never changed at all; rather, as can happen in other sorts of judgments and observations, people simply misjudged that this person was an X. Through re-evaluation, we might, indeed, find that X’s simply do not exist and that the judgment “He is an X” is false and was always false. It is always possible to misjudge the facts.

In further lines, Hare goes on to establish a special dictum to represent his position. “‘No substantial disagreement without verbal agreement,’” he states, “is a useful slogan... which is fatal to naturalism and to most other kinds of descriptivism.”²⁸ His attempt here, and throughout *Moral Thinking*, is to demonstrate that, because moral disagreement occurs, moral words are dependent solely

²⁶ Hare, *Moral Thinking*, 69.

²⁷ Oderberg, *Real Essentialism*.

²⁸ *Ibid.*, 69.

on the standards of each individual and not on any naturalistic meaning. Morality, of course, is the only context in which attributive adjectives are often used and in which verbal disagreement occurs. Such words as “fragile,” “wide,” “long,” and “light” are never used attributively and can never incur disagreement. This is all facetious, of course. I might have some differing standard of width than you if I see some desk which I have never seen before. I might assert “That is a really wide desk!” Another person might, additionally, come along and assert “No, that is not a wide desk at all. It is actually quite narrow.” Yet another person might come along and assert “Well, I think it is somewhat wide, but not too wide.” Replace the word “wide” with the word “good,” the word “narrow” with the word “bad,” and the word “desk” with, say, “nuclear weapon” in all of these sentences, and you have, depending on the circumstances, moral disagreement. It is not the case that, simply because I disagree with how someone is using an attributive adjective, I misunderstand the meaning of the adjective in question.

Moral disagreement, of course, can occur for a multiplicity of other reasons as well. Morality concerns values, and, suffice it to say, we care very much about values. People can, and certainly do, establish unsound standards for themselves. To challenge or deconstruct these standards can often have unwanted psychological consequences for her whose standards are being thus treated. It is no wonder, then, that people often become very defensive about their values, even if they are unsound ones; and, often, no amount of argument will suffice to result in a rejection of those values.

Having considered the obfuscatory nature of such words as “commendation,” “standard,” and “value-word,” as well as the preceding arguments against Hare’s account of descriptive meaning and evaluative meaning, my elucidation of what might well be called the non-descriptivist fallacy, and my arguments on predicative and attributive adjectives, we should be able to see more clearly the plethora of problems in Hare’s following comments in *Sorting Out Ethics*:

Although evaluative (including moral) statements do indeed have truth conditions, these can change without the entire meaning of the sentences which express them changing. This has crucial consequences for ethical theory. If we change the truth conditions of a moral statement, we change its descriptive meaning. But if the evaluative meaning remains the same, we have, in making this change, altered our moral standards. We are appealing to different reasons, for example, for calling an act wrong, but we are calling it wrong in the same sense, evaluatively speaking. We are still condemning it by calling it wrong.²⁹

Having concluded that non-descriptivism is a fundamentally unsound framework and that, therefore, some form of descriptivism must be sound, we have good reason to attempt to discover a sound descriptivist model. We may take this even further and, eschewing non-naturalism for its sheer ambiguity and tendency toward intuitionism, which is fundamentally unsound, make the choice to attempt to discover a specifically naturalist model, especially since I, and many others, have laid some solid groundwork for demonstrating its viability. Our main concerns toward naturalism

²⁹ Hare, *Sorting Out Ethics*, 54.

should be focused, then, on discovering the natural property or properties which define “good.” The bulk of this work will be discussed in later sections. For now, let us examine Hare’s ethical theory, itself, and determine its merits, demerits, and its ultimate relationship to the justification of ethical judgments.

3.1.2 *Hare’s Dialectical Method*

Hare’s universal prescriptivist account of ethics has been ground-breaking. He has at once achieved two feats: he has both greatly supplemented the work of other theorists in refuting subjectivism and relativism, and he has shown, via his hypothetico-deductive method, that ethical judgments can and do follow logical rules. Gewirth and Beyleveld have, I think, rightly characterized Hare’s theory as a dialectical one, in that the soundness of ethical syllogisms is dependent on the statements that agents make. Hare’s method is, however, dialectically contingent; the soundness of the practical syllogisms within his theory rests upon the statements that agents *incidentally* make. This is contrasted with a dialectically necessary method, in which the soundness of the practical syllogisms rests upon the statements that agents *must* make insofar as they are agents. Both methods yield feasible judgments, but the true test is whether they yield *justifiable* judgments. I shall now explore whether this is the case with regards to Hare’s theory.

In *The Language of Morals*, Hare establishes the groundwork for his theory. As such, he focuses his attention primarily toward the structure of practical principles. Prior to or during our voluntary actions, Hare asserts, we make judgments. These judgments need not be moral; they might be prudential or focused toward specific activities such as driving. All such judgments, he maintains, yield principles of action. Such judgments can be represented by a simple practical syllogism. The major premise concerns a proposed mode of conduct, and the minor premise concerns what one should do given some alternative or set of alternatives. “The two premises,” he states, “create a principle.”³⁰ Such judgments, he claims, necessarily concern specific criteria relevant to a particular situation such that, if one were ever in a relevantly similar situation constituted by the same criteria, one’s originally formed principle would necessarily apply to that situation as well. Our judgments, therefore, contain an element of universality and, thus, so do our principles. For such activities as driving, swimming, walking, and so on, we, at some point, learned to apply universal principles so that we could participate in these activities in a variety of contexts. As Hare affirms: “Everything we are taught must be reducible to principles—taught either by example or by precept.”³¹ For ethical judgments, Hare affirms, there is no reason to expect that this basic pattern should be any different. Such judgments, in order to have prescriptive force, he

³⁰ Hare, *The Language of Morals*, 56.

³¹ *Ibid.*, 159.

claims, must express commands; in order to fulfill this criterion, an imperative must comprise one of the premises in the practical syllogism in question. As he asserts: “No imperative conclusion can be validly drawn from a set of premisses which does not contain at least one imperative.”³² It is at this point that a vital part of Hare’s thesis becomes clear: “A command, like a statement, must observe certain logical rules,” namely that of non-contradiction, by virtue of “the presence of logical connectives in the phrastics” of imperative statements.³³ Thus, on Hare’s account, if one assents to an imperative in a particular instance, one implicitly accepts the general principle derived from the relevant criteria upon which the formation of the imperative is based. When one forms an imperative judgment, therefore, one must accept a universalized form of the imperative which applies in all relevantly similar circumstances, where the relevant similarities are contained in the principle.

In *Freedom and Reason*, he builds from his general groundwork and gives greater clarity and form to his theory. One of the first issues he addresses is his thesis of universalizability. “If a person says ‘I ought to act in a certain way, but nobody else ought to act in that way in relevantly similar circumstances’,” he states, “then, on my thesis, he is abusing the word ‘ought’; he is implicitly contradicting himself.” Hare’s main purpose, here, is to demonstrate that consistency in one’s practical judgments is key: if I utter “B,” then I cannot utter “¬B” without inconsistency, where B represents the relevant criteria cited in my initial prescription. Hare uses a scenario to illustrate his theory in action. He asks us to consider a situation in which some person A owes some person B money and that B owes some person C money, but A cannot repay B, and B cannot repay C. B makes the judgment that A ought to be put in prison for his inability to pay. The results, Hare explains, can be summed up as follows:

B asks himself, ‘Can I say that I ought to take this measure against *A* in order to make him pay?’ He is no doubt *inclined* to do this, or *wants* to do it. Therefore, if there were no question of universalizing his prescriptions, he would assent readily to the *singular* prescription ‘Let me put *A* into prison’. But when he seeks to turn his prescription into a moral judgement, and says, ‘I *ought* to put *A* into prison because he will not pay me what he owes’, he reflects that this would involve accepting the principle ‘Anyone who is in my position ought to put his debtor into prison if he does not pay’. But then he reflects that *C* is in the same position of unpaid creditor with regard to himself (*B*), and that the cases are otherwise identical; and that if anyone in this position ought to put his debtors into prison, then so ought *C* to put him (*B*) into prison. And to accept the moral prescription ‘*C* ought to put me into prison’ would commit him (since, as we have seen, he must be using the word ‘ought’ prescriptively) to accepting the singular prescription ‘Let *C* put me into prison’; and this he is not ready to accept.³⁴

From this example, we begin to observe what is really going on in Hare’s theory. The emphasis, as he plainly acknowledges elsewhere, is on personal preferences. This is well-captured in his clarification of the preceding example:

³² *Ibid.*, 28.

³³ *Ibid.*, 24.

³⁴ Hare, *Freedom and Reason*, 91.

In the example which we have been using, the position was deliberately made simpler by supposing that *B* actually stood to some other person in exactly the same relation as *A* does to him. Such cases are unlikely to arise in practice. But it is not necessary for the force of the argument that *B* should *in fact* stand in this relation to anyone; it is sufficient that he should consider hypothetically such a case, and see what would be the consequences in it of those moral principles between whose acceptance and rejection he has to decide.³⁵

The principles upon which one must make one's decisions, on Hare's account, are characterized by a weighing of preferences in accordance with personal principles. Herein lies the trouble: Hare has not set forth any definitive reasons why our personal preferences constitute, in the first place, justifiable sources of moral judgment. If, as Hare strongly implies throughout his works, "preferences" implies "likes" and "desires," then we need only re-examine the structure of ethical judgment to discover that preferences and likes do not constitute the basis of all such judgments. It is, first, vital to clarify exactly what I mean. I do not wish to suggest that there exists some voluntary action that is unrelated to desire, for, certainly, in order to make a practical judgment in the first place, one must, at least in some very loose sense, desire the object or state of affairs contained in one's judgment. Hare's use of "preferences" and "likes," however, seems to imply something different. "Preferences," as he uses the term, seems to imply something along the lines of "that which pleases." If this is the nature of his usage, then there is something horribly wrong with his account at the outset, for we are stuck back in the old hedonist predicament: "Why should I make judgments or form principles based upon what pleases me or others?" To make my point even clearer, Hare seems to be referring, specifically, to first-order desires through his use of the word "preferences," and this failure to distinguish between first-order and second-order desires, the latter which constitute desires about first-order desires, is problematic for the following reason: there is no principle of coherence in Hare's theory, besides his ambiguous notion of "intensity," to instruct us as to which of the two categories of desire we are to draw from in order to form our practical judgments. If *X* has a first-order desire for some object or state of affairs *Z* and *Y* has a second-order desire for some object or state of affairs *Z*, yet *X*'s first-order desires are "more intense" than *Y*'s second-order desires, and there is a conflict between *X* and *Y* as to who should obtain *Z*, then how are we to prescribe who should obtain *Z*? Although Hare addresses relevantly similar scenarios, it is telling that he either never quite approaches a solution or that he calls the situation unrealistic and tweaks it to get the result he wants. These types of scenarios, however, are quite realistic indeed. R. B. Brandt hints at such a problem in his "Act-Utilitarianism and Metaethics." Citing one of the scenarios Hare offers in *Moral Thinking*, he writes:

Suppose I come to a parking space in my Mercedes-Benz 450 SL, currently occupied by a bicycle. The space is the only vacant one in sight, near to a grocer's shop in which I wish to shop. Is it wrong for me to move the bike, standing it against a tree, so that I can park my car? Hare says, 'No'—*provided my rational and prudent experience-based desire to park my car is greater than the bike-owner's similar desire that his bike not be moved.* For I can

³⁵ *Ibid.*, 93.

universally prescribe moving the bike only if, given these packages between which moral language requires me to choose, moving-bike-plus-parking-car and not-moving-bike-and-no-parking for all possible worlds (including one of me having my bike moved if I am the bike-owner), I prefer the former. This is also the conclusion which the act-utilitarian will reach provided the above experience-based preferences are rational and prudent.³⁶ (*italics mine*)

Brandt's point here is that the rightness of the act in question is highly subject to the agent's and patient's respective psychological states and predispositions. The rightness of the consequences, on Hare's theory, will depend on the intensity of each party's respective preferences. In his response to Brandt,³⁷ he acknowledges that "causal factors" might impede our ability to represent to ourselves what it will be like to be in a given situation but still insists that, in order for us to make the right decision in such a situation, we must have preferences of intensity equal to those of our patient. He seems, then, to overlook a major part of Brandt's objection, namely that such preferences will not always be of equal intensity. Consider the consequences of this. On this thesis, an agent may do anything she wants to any patient, as long as the patient prefers not to be acted upon in this way by the agent less intensely than the agent prefers to act in this way. For instance, if someone attacks a person who has only a mild preference not to be attacked, and this is at least conceivable, then to attack would be a permissible decision. One may replace the verb "attack" with any verb, if one is not satisfied with my example. There is, thus, a lack of coherence concerning which preferences, whether first-order or second-order, have priority. Apparently, it does not matter; Hare's theory seems, thus, to take on a Benthamite hue. Even if Hare imported some principle of coherence into his theory to rank-order desires by some scale other than intensity, as, indeed, John Stuart Mill did, he would still face the hurdle of justifying, as it were, internalist reasoning.

Internalist reasoning, in general, is a theoretical model which acknowledges no reasons for action other than those that are internal to the agent. There is a meaningful distinction to be made between types of internalism with regards to motivation as a normative justificans. Some accounts of internalism conceive of an agent's desires, whether first- or second-order, as constitutive of the exhaustive criteria available for ethical justification. Others, however, conceive of internalism in significantly different terms. Nagel, for instance, characterizes internalism as more generically motivational, while he rejects purely desiderative conceptions. Smith characterizes internalism in a similar yet slightly different manner such that mere internal desiderative reasons would not count as a justificans for action.³⁸ On his account, rather, it is belief-motivated desires concerning what is objectively

³⁶ R. B. Brandt, "Act-Utilitarianism and Metaethics." *Hare and Critics: Essays on Moral Thinking, with Comments by R. M. Hare*, Ed. Douglas Seanor and N. Fotion. (Oxford: Oxford University Press. 1988): 229.

³⁷ Douglas Seanor and N. Fotion, *Hare and Critics: Essays on Moral Thinking, with Comments by R. M. Hare*, (Oxford: Oxford University Press. 1988): 216.

³⁸ Michael Smith, "Internal Reasons," *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics*. Ed. Michael Smith. (Cambridge: Cambridge University Press, 2004): 17-42; "Internalism's Wheel," *Ethics and the A Priori*, 318-342.

desirable that constitute such a justificans. When I speak of internalism with regards to ethical justification, however, I will refer to the desiderative model. This desiderative-internalist approach to justification is one I wish to refute.

Hallvard Lillehammer, Hampton, and others have exposed internalist models as deeply flawed. Externalists, by contrast, claim that agents have reasons to act that exist independently of their desires. “Their claim,” Lillehammer explains, “is that reasons which derive from a rationally sound development of the present desires of an agent are not the only reasons for action there could be. Agents might have reasons which derive from the desires of an agent at different times, from the desires of a different deliberator, or perhaps even reasons which exist independently of practical reasoning as such.”³⁹ However, he explains, externalists have, thus far, failed to deliver definitive arguments demonstrating the flaws in internalist models. Lillehammer sets out to deliver such an argument. His central thesis, he explains, is that “the doctrine of internal reasons fails on its own terms as a direct consequence of its commitment to the notion of a rationally sound deliberative route.” As he clarifies: “If the existence of a rationally sound deliberative route entails the existence of rational requirements on action which are not constrained by the contents of the present desires of agents, it follows that reasons which depend on the upshot of a rationally sound deliberative route entail the existence of external reasons, thereby contradicting the doctrine of internal reasons.”⁴⁰ Internal reasoning, then, he claims, rests on self-defeating premises.

One of the principal claims upon which the soundness of the internalist account hinges is that internalist reasoning is a sound deliberative method of reasoning. Lillehammer, however, rightly states that, if this is what internalists wish to claim, they must provide “an account of what makes a deliberative route rationally sound, and which retains an interesting link between reasons and the present desires of an agent.”⁴¹ The internalist, he explains, has two options. As he states:

Either what counts as a rationally sound deliberative route for an agent is relative to the content of his present desires, or it is not. Let us call the first option process-relative and the second option outcome-relative. For to process-relative internalism, the reasons of an agent are determined by the upshots of a rationally sound deliberative route, where both the outcome of the route and the nature of the route are constrained by the present desires of the agent. For outcome-relative internalism, the reasons of an agent are constrained by the upshots of a rationally sound deliberative route, but the nature of the route itself is not constrained by the content of the present desires of the agent.⁴²

He goes on to argue that both the process-relative and outcome-relative models fail. He begins by considering the process-relative model. There are three approaches, he explains, that an internalist can take under this model. As he states:

³⁹ Hallvard Lillehammer, “The Doctrine of Internal Reasons,” *The Journal of Value Inquiry*. 34, no. 4, (2000): 508.

⁴⁰ *Ibid.*

⁴¹ *Ibid.*, 510.

⁴² *Ibid.*

First, the internalist can define a rationally sound deliberative route in terms of a process which an agent presently desires to engage in, where this entails that the agent has a desire with a content to the effect that he will reason this way. Call this conception principles as ends. Second, he can define a rationally sound deliberative route in terms of a process the employment of which would promote the agent's present desires. Call this conception principles as means. Third, he can define a rationally sound deliberative route in terms of a process the agent would either desire to engage in, or which would promote the satisfaction of his desires, if he were practically rational. Since an internalist defines practical rationality in terms of the upshots of a rationally sound deliberative route, this strategy boils down to the claim that what counts as rationally sound deliberation is itself determined by the upshot of rationally sound deliberation.⁴³

Ultimately, all three accounts are vacuous and explain very little, if anything. Lillehammer examines each approach and concludes that process-relative internalism is "intrinsically implausible." Ultimately, he explains, both accounts share a common problem: "The crucial question of what makes a deliberative route rationally sound remains unanswered, thereby threatening the internalist position with emptiness, regress, or circularity..." and, thus, even the outcome-relative interpretation fails to make its point, for outcome-relative internalism, in fact, "collapses into externalism."

Hampton argues along relevantly similar lines, but she additionally emphasizes the normative mistakes inherent in internalism. The internalist, she explains, "must treat reasons to deliberate as 'external' in just the way that his opponents do—that is, these reasons must be regarded as external *in the sense that they apply to, and rightly direct, an agent regardless of whether she likes the idea of their doing so or not.*" (italics mine) "It seems," she adds, "that he cannot object to the idea of there being external reasons, because his own account assumes them, insofar as it assumes that we have reasons to deliberate in certain ways, at certain times, no matter our existing motives."⁴⁴ Thus, the internalist account not only relies on the positing of external reasons, but it also normatively prescribes such reasons. We should thus realize the circularity inherent in any theory which is founded on premises concerning an agent's desires or preferences.

Hare's theory, therefore, makes two mistakes at the outset: not only does he fail to provide us with a reason why we should adhere to prescriptions concerning preferences, but he also fails to establish any principles of coherence to allow us to discover which *types* of preferences, and not simply which preferences, to accept as legitimate or justifiable ones. It is most likely Hare's non-descriptivism which renders the former flaw. Certainly, I will be bound to accept the truth of "P(a)," as an instance of the universal statement " $\forall x(P(x))$," but I will only be so bound if the statement " $\forall x(P(x))$ " is true, and proving " $\forall x(P(x))$ " to be true requires that "P(a)" not be false. The trouble is that Hare has not shown "P(a)," namely, "I ought to do X" and the suppressed premise "I ought to base my practical judgments upon preferences," to be not-false. The Harean might object that proving "P(a)" to be not-false is not

⁴³ Ibid.

⁴⁴ Hampton, *The Authority of Reason*, 78.

meaningful and that we must simply prove that “P(a)” is formally consistent with “ $\forall x (P(x))$,” that is, that our prescriptions are formally valid, since soundness cannot apply to singular moral judgments. This, however, is an assumption and not a proof. Thus, because of its non-descriptivist assumptions, we see that the Harean position is simply valid and not sound. Furthermore, Hare implicitly admits that his own theory is not sound, and this on the false assumption that soundness is not necessary.

Let us, however, for the sake of argument, assume that preferences are sound bases for moral judgment. Keeping this assumption in mind, let us revisit the basic structure of the account. We must, Hare states, establish some principle to navigate a specific situation. In order to establish such a principle, we must consider whatever we take to be the relevant criteria of that situation. Once we have established what we interpret to be the relevant criteria of the scenario, we must make a judgment in the form of the imperative or ought statement “I ought to do action X to patient Y in situation Z.” In *Moral Thinking*, Hare presents the following scenario to illustrate this process:

A clear example of such a relevant feature would be the fact that, if I now drove at an interval of less than so many cars’ length from the vehicle in front (given the speed at which we are travelling), and it had to stop abruptly, we should have a collision and several people, dogs, etc., would be hurt. The candidacy of such a feature will be amply supported and confirmed in what follows, when we discuss how we are to decide what moral principles to accept. We shall see that the method of critical thinking which is imposed on us by the logical properties of the moral concepts requires us to pay attention to the satisfaction of the preferences of people (because moral judgements are prescriptive, and to have a preference is to accept a prescription); and to pay attention equally to the equal preferences of all those affected (because moral principles have to be universal and therefore cannot pick out individuals).⁴⁵

I can choose whatever criteria I want to be relevant, but, Hare warns, I better choose wisely. If I establish the prescription “I ought to crash into and injure the person in front of me while driving on this road,” I, thus, establish the principle “P ought to crash into and injure persons in front of P while driving on the road.” On Hare’s account, it follows from this that “P ought to crash into and injure me (if I am in front of P) while driving on the road.” If we consider such a prescription, Hare asserts, we will not be willing to universally prescribe it, since we would not prefer to be treated thus in such circumstances. There are two problems, however, which arise from such a formulation. The first is that there are major inconsistencies in Hare’s collective accounts of how one’s preferences are to be applied. There seems to be dissonance in his theory over whether, in such a situation, I am to apply my own present preferences to my patient’s preferences and experiences or whether I am to hypothetically adopt my patient’s preferences. As Hare correctly states, there is a significant difference between the propositions “I now prefer with strength S that if I were in that situation x should happen rather than not” and “If I were in that situation, I would prefer with strength S that x should happen rather than not.”⁴⁶

⁴⁵ Hare, *Moral Thinking*, 92.

⁴⁶ *Ibid.*, 95.

The former proposition concerns my own present preferences, and the latter proposition concerns the patient's preferences or, in other words, the preferences that I would have in such a situation if I had the preferences of the patient. Both readings yield significant problems for Hare's theory. The second problem is that, despite Hare's suggestion that singular prescriptions cannot be universalized, they, indeed, can be; they will simply be vacuously universal.

We shall tackle the former issue first. In *Moral Thinking*, Hare gives us the following instructions for applying preferences:

I emphasize that the imagined situation must be one in which I have [*the patient's*] preferences. If, by some quirk of nature, I were a person who knew that he did not feel pain in that situation, or if I knew that I was going to become such a person by being anaesthetized, then I might indeed sincerely say that I did not mind being subjected to the experience. . . . But this would be irrelevant; and so would it be if I knew that I would feel pain, but for some reason would not mind it. For I am to imagine myself in his situation with *his* preferences. Unless I have an equal aversion to myself suffering, forthwith, what he is suffering or going to suffer, I cannot really be knowing, or even believing, that being in his situation with his preferences will be like *that*.⁴⁷

Even this account seems conflicted. Am I to hypothetically *consider* his preferences as relevant even if I cannot *know* his preferences, or am I merely to know his preferences and then consider my own preferences hypothetically applied to his situation? Hare later attempts to clarify his position by telling us that "If I now say that I ought to do a certain thing to a certain person, I am committed to the view that the very same thing ought to be done to me, were I in exactly his situation, including *having the same personal characteristics and in particular the same motivational states*."⁴⁸ (italics mine) Here, therefore, he seems to make it clear: we must consider ourselves *exactly* as our patient, with the same exact preferences of our patient. In Hare's reply to A. Gibbard's objections to his theory, Hare tells us, "A conditional preference is a preference I actually have for a hypothetical circumstance. It may not be a preference I *would* have if I were in that circumstance.' Here Gibbard, unlike many writers, gets this important point right."⁴⁹ With all of these conflicting accounts, we are left befuddled as to how we are to apply preferences to the situations with which we are faced, and this is quite troubling.

Let us examine the consequences of both interpretations. First, let us consider the hypothetical application of our patient's preferences. On this account, we must form a prescription for which, if we were to act on it, we would prefer to have the act in question done to us if we were in the patient's position with her preferences. Let us examine the logical structure of this more closely. Again, we must start with the following proposition: "I ought to do action X to patient Y in situation Z," with the addition of the conjunct "and patient Y has patient Y's preferences." On Hare's account, therefore, it follows, by universalization, that "P ought to do action X to

⁴⁷ Ibid., 94.

⁴⁸ Ibid., 108.

⁴⁹ Seanor and Fotion, *Hare and Critics*, 229.

patient Y in situation Z, and patient Y has patient Y's preferences." It follows from this that "P ought to do action X to *me* in situation Z, and I have my original patient's preferences." But does this really follow? Does it not instead follow from "P ought to do action X to patient Y in situation Z, and patient Y has patient Y's preferences" that "P ought to do action X to *me* in situation Z, and I have *my own preferences*?" Where, on this interpretation, could we possibly logically derive that we must hypothetically imagine ourselves to have our patient's preferences? It seems as though Hare is asking us to seriously consider the statements "If I had my patient's preferences in situation Z, then I would have my patient's preferences in situation Z" or "If I had the preferences of my patient not to be acted upon in manner X in situation Z, then I would not prefer to be acted upon in manner X in situation Z," as if these were not tautologies. Even if this could be sorted out logically, Hare would face the further obstacle of demonstrating to us why we should include in our criteria of relevant similarities our patient's preferences in situation Z if we could never, or even will never, have such preferences in situation Z. On this point, there seems to be a great degree of logical incommensurability, for how is it even intelligible for an agent to include in a maxim those criteria whose universalized contents will never apply to him? In *Sorting Out Ethics*, Hare seems to consider this possibility, but shrugs it off without much argument:

There is a further problem about whether being actual as opposed to merely possible or hypothetical is a universal property. If it were, a form of special pleading would become possible in moral reasoning, by which an aggressor could claim that he would never be actually in the position of his victim, and that this difference was morally relevant. It is perhaps best to follow those who claim that the actual world cannot be distinguished from possible worlds without a reference to individuals, namely those who are actual; but not to follow them into thinking that possible worlds have some real existence in limbo.

This problem of incommensurability, however, is a very serious one for Hare's theory. If I can never have the specific preferences of my patient, can I ever really know what it is to have her preferences? If I could never have her preferences, and if, thus, I would never be in her circumstances in the first place, then how could such preferences ever be considered relevant, on any interpretation of this word, to a maxim concerning my preferences? This leads directly to the second interpretation.

On the second interpretation of Hare's account of universalization, we must consider our own preferences as relevant criteria. This account is far more sensible but has some serious flaws; it does not take into account situations in which people might actually have different or conflicting preferences in relevantly similar situations. Yet again, we must formulate the generic prescription "I ought to do X to patient Y in situation Z." On Hare's account, it follows from this that "P ought to do X to patient Y in situation Z." Hare's use of the car-accident example is quite fitting for this kind of universalization, but this is only because both the agent and the patient in this situation share the *same* preferences, namely not suffering and not being injured. On the assumption that his method of universalization is valid, Hare rightly states that, if one prescribes that one ought to injure one's patient, then one must prefer the consequences of one's actions more intensely than one would not want to be so acted upon. These sorts of people, what Hare terms "fanatics," are

rare; on this basis, he states, we need not worry about them. Leaving aside this problematic feature, let us consider a situation in which the agent and the patient's preferences are incommensurable. Consider that I am a racist senator. Because I am racist, I consider all interracial marriage to be abhorrent; I have an intense preference directed against it. Thus, I prescribe "I ought to enact legislation that will abolish interracial marriage." On Hare's theory, it seems that this prescription would be justifiable, for I will never have a preference to be part of an interracial marriage. One could say, and people do say, the same about same-sex marriage and polygamy. One might make the objection that it is the fact that I am forcibly preventing people from doing what they prefer that is morally relevant. On Hare's theory, however, this is precisely not morally relevant, since I determine what is morally relevant by what I include in my maxim and not by what is *implicit* in my maxim; furthermore, even if we were to follow the rule that we must take into account what is implicit in our maxim, we could only consider those implicit things morally relevant that constitute *commonly held* preferences between the parties, since, as we have seen, some preferences are simply incommensurable on Hare's account.

Thus, on both interpretations of Hare's method of preference application, we must conclude that, if we assume that preferences are sound bases for moral judgment, that his method of preference application demands preference commensurability between agent and patient; that is, the preferences must be *commonly held* between the parties involved. This leads us, indirectly, to the second main issue with Hare's theory: that of his method of universalizability.

The reader might have noticed that, throughout my discussion on Hare's preference-application problem, I repeatedly emphasized that we must assume that his method of universalization is valid. This was purposeful. Let us now analyze this issue more closely. Once again, the generic form of our prescriptions is "I ought to do X to patient Y in situation Z." The fundamental step that follows, a step that is fundamental to Hare's theory, is the inference "P ought to do X to patient Y in situation Z." This inference is fundamental precisely because this is the claim, laid out in *The Language of Morals*, that provides the groundwork for the rest of his theory. To refresh a bit, his claim is that, when an agent makes a practical judgment, she forms a general principle for her action so that, in relevantly similar situations, she can apply the principle to direct her actions. This does not seem to me to be a wild claim on Hare's part, but let us examine its implications. The first problem is that, although one might accept some principle for one's action, one might be incredibly indecisive and, in a relevantly similar situation, might choose to adopt some other principle. In other words, I can change my mind. Certainly, any ethical theory worth its salt will allow me to change my mind about or re-evaluate an action; once I have decided on some general principle, must I be bound to it forever? The answer is no. Suppose I am driving my car down the road; I have established some principle or principles of driving. These principles include "Drive with your hands on the wheel," "Brake to stop the car," and "Step on the gas pedal to move the car." I need not, however, accept these principles for every situation. I might be a very reckless driver and decide to drive with my feet, or I might decide

that, as dangerous as it is, I will only step on the gas pedal and never on the brake. Keep in mind that I am citing logical exceptions to Hare's thesis; regardless of what Hare says about the likelihood of specific scenarios or decisions, it is *logical* counterexamples that matter.

I might also form principles in another way. Instead of changing my principles constantly, I might choose to form extremely specific principles such as "I ought to do action X that has properties $\Phi 1, \Phi 2, \Phi 3, \Phi 4,$ and $\Phi 5$ to patient Y if she is $\Psi 1, \Psi 2, \Psi 3, \Psi 4, \Psi 5,$ and $\Psi 6$ in situation Z that has properties $\delta 1, \delta 2, \delta 3, \delta 4, \delta 5, \delta 6, \delta 7, \delta 8,$ and $\delta 9.$ " If I am almost obsessively thorough in my formulation of principles, I can always formulate them such that they could never contradict one another, although other individuals might consider the small details that I pick out to be irrelevant.

Let us consider a related problem with his method of universalization. This one concerns Hare's following claim in *Moral Thinking*:

Given two cases differing solely in that in one of them individuals A and B occupy certain roles, and in the other the roles reversed, any universal principle must yield the same prescriptions about them in both. In order to yield different prescriptions about the two cases, the principle would have to contain the names of the individuals, and would therefore not be universal.⁵⁰

But is this really correct? Is it the case that a prescription is not universalizable if an agent cites a name as a relevant criterion? Let us scrutinize the generic form of Harean prescriptions. In the prescription "I ought to do X to patient Y in situation Z," I might claim myself and some other specific person as relevant criteria by reformulating the statement thus: "I (Stephen) ought to kill James by running him over." Universalization, then, yields the following statement: "All people that are me ought to kill all people that are James by running them over." Hare might, once again, object that this is not really universalization. On the contrary, it is, indeed, universalization. Simply because all instances of me are me and all instances of James are James, it does not follow that, formally, this is not a case of universalization. It might be vacuous universalization, but it is universalization all the same. A Harean might object that I have universalized this the wrong way. I must, he might insist, imagine myself as James with James's preferences. Considering my previous observation that it is commensurable, *commonly held* preferences between parties that ultimately give Hare's theory some semblance of justificatory weight, and also considering that killing is antithetical to living and that living is usually a commonly held value between parties, we might consider this a biting objection. Let us, then, explore this. According to this account, and still citing only names as relevant, we must prescribe the following: "If I am James, James ought to kill me by running me over." This, however, is highly problematic, for I can never be James; this would violate the law of identity. Therefore, to begin with, the statement is altogether implausible, for it entails the following prescription: "If I am James, James, who is also me, ought to kill me by running me over."

⁵⁰ Hare, *Moral Thinking*, 112.

Thus, Hare's claims concerning the wide circumstantial applicability of his theory are doubtful. The main problems can be summed up as follows:

1. **The Preference-Legitimacy Problem:** The theory suffers from the task of demonstrating that preferences are sound bases for moral judgment. In fact, anti-internalist arguments give us good reason to reject such a basis for moral judgment.
2. **The Preference-Coherence Problem:** Even if we assume that preferences are sound bases for moral judgment, the theory still suffers from incoherence in the ranking of preferences by relative intensity and first-order versus second-order desires.
3. **The Preference-Incommensurability Problem:** Even if we assume that relative intensity of preferences is the final arbiter in our treatment of our patients, the theory still suffers from a lack of applicability to all situations due to the possibility of incommensurability of preferences between parties.
4. **The Vacuous-Universalization Problem:** Even if we assume that the problems in 1–3 are not problems, the theory still suffers from a problematic method of universalization that leaves open the possibility of vacuous universalization. The possibility of vacuous universalization leaves open the possibility that agents may act in whatever ways they wish, and this results in a theory that is no different, and thus also no more interesting, than a theory of ethical egoism.

Thus, in order to even begin to consider Hare's thesis, we must jump over four huge philosophical hurdles, or we must simply shrug off these problems and assume that all is sound. Neither of these choices are good ones. There is, however, a silver lining in Hare's theory. Recall the Preference-Incommensurability Problem; this problem is only a problem for some preferences and not all of them. It is certainly possible for people to have various, idiosyncratic values but, as Hauser and a number of other anthropologists, biologists, and sociologists have slowly come to discover, there is much cross-communal and cross-cultural consistency among certain values. Among universal disvalues, Hauser includes "killing, causing pain, stealing, cheating, lying, breaking promises, and committing adultery." He qualifies, "Like other rules, these moral rules have exceptions. Thus, killing is generally forbidden in all cultures, but most if not all cultures recognize conditions in which killing is permitted or might be justifiable. . ."⁵¹ Thus, there is reason, he concludes, "to believe that there are universal properties of the human mind that constrain the range of cultural variation."⁵² However universal, however, it does not follow from our having such a universal moral grammar that this universal "moral faculty," as Hauser terms it, gives us justifying reasons for action. Given the general validity of Hare's theory, however, it *does* follow from a collaboration with Universal Moral Grammar Theory that the presence of such universal values provides us with some solid grounds for a negative theory of ethical

⁵¹ Marc D. Hauser, *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*, (New York: HarperCollins Publishers. 2006): 48.

⁵² *Ibid.*, 419.

justification. Thus, while I have not, as of yet, presented the framework through which we can make justifiable ethical judgments, it can be shown that Universal Prescriptivism, aided by Universal Moral Grammar Theory, yields a basis for picking out unsound attempts at ethical justification. Apart, it seems, the two theories lack justification but, when combined, they form a vital part of a negative theory of ethical justification. Let us explore this.

3.2 Developing a Method of Justification

In this section, I develop an account of negative value justification. I argue that this is a necessary step because, before we can demonstrate that it is justifiable even to go about developing a positive theory of value and discarding such theories as utilitarianism, egoism, relativism, and skepticism, we must show that these theories have a certain burden of proof, on their own grounds, which my account does not have. As such, I methodically demonstrate how a burden of proof in value theory can be determined and, additionally, how we can pragmatically and deductively show that certain value statements are false. This proves to be an important methodological stepping stone toward our exploration of positive dialectical theory. Thus, the methodological goals of the negative justificatory approach are twofold: (1) to prove, via the demonstration of the existence of necessarily false value judgments, that value judgments are subject to truth conditions and thus that it is *prima facie* justifiable, in the first place, to go about developing a normative theory of value and (2) to prove, via the instantiation of necessarily false value judgments, that certain other theories about normative value are false and thus that it is *prima facie* justifiable to go about attempting to develop a specific kind of normative theory of value.

3.2.1 *The Problem of a Justificatory Method*

Imagine you are walking along and you see someone walking in front of you. Suppose, also, that you are feeling friendly; you engage the person. Also feeling friendly, she reciprocates. You begin to delve into an interesting conversation and, eventually, out of the blue, she points to someone in the distance and tells you that she is going to attack him and steal his money. Having conversed with her for some time, you realize that she is amenable to tempered discussion and persuasion, and so you decide to engage her on this issue. “But you shouldn’t do it,” you tell her. She then asks, “Why?” How are you to defend your position? Of course, you might respond “Because you would make him unduly suffer and deprive him of his property,” or “Because you would harm him,” or “Because it would violate his human dignity.” The question “Why?,” however, persists seemingly *ad infinitum*. Another question that so persists is “Why should I care?” Thus, in ethics, we begin

to see the problem of justification. A Hobbesian might tell us that it does not make sense for her to attack and rob the other person because she has an interest in keeping consistent to the rule “Do not attack and rob,” lest she some day be attacked and robbed. A utilitarian might tell us that such an action is against the principle to maximize utility. A Humean emotivist might tell us that, after having considered all of the criteria relevant to the situation, most people would feel bad for the would-be victim and would reject the attacker’s actions. The accounts go on and on, but most of them suffer from a fatal flaw: they do not provide a justification that the attacker could not validly rebut. To the Hobbesian, she might respond, “But if I attack and terrorize this person and, indeed, enough people, I might gain enough power such that it will be in my interest to keep attacking people. Why can’t *I* be the Leviathan? If I have enough power, I do not need rules.” To the utilitarian, she might simply say, as did Nietzsche, that she does not care about pleasure or, at least, other people’s pleasure.⁵³ “Pleasure,” she might say, “is the privilege of the strong.” She might also tell you that she seeks to be, in Nozick’s terminology, a “utility monster,” depriving everyone in society of his pleasure in order to maximize her own, which is perfectly consistent with many forms of utilitarianism.⁵⁴ To the Humean emotivist, she might tell you that she either does not care about emotions or that she cares only about her own emotions. She might also point out the fallacy of appeal to the majority inherent in Hume’s theory.

There is an implicit, repeated theme in her reasoning, and that is “I don’t care.” This is a vital point, and it is one to which we must respond. The key question here is: how can you prove to her that she *should* care? What can you offer to demonstrate to her that she *ought not* attack and rob her would-be victim? “Ought,” it seems, is at the center of any discussion on ethics, and so, therefore, is a serious analysis of statements containing “ought.” I will not undertake such an analysis here, although I will do so in Chap. 4. For now, I will simply make an important point and build from there: in order to justify any ethical statement, one must prove that the statement is true. Neither Universal Moral Grammar Theory, nor utilitarianism, nor emotivism, nor intuitionism, nor contractarianism, nor many forms of deontology prove that such statements as “I ought to do X” and “X is good” are true. How do we know this? There are counterexamples, ones that often fit within the standards of the theories themselves, and no matter how unlikely the counterexamples are, no matter how much they seem contrary to the interests of the agent, they are counterexamples nonetheless. How, then, can we meet her challenge? For one, we can begin by building a *negative* method of justification, that is, we can examine and demonstrate what sorts of statements *do not* constitute ethical justifications. We can, in other words, begin, not by demonstrating which kinds of ethical statements are true but, rather, which kinds of ethical statements are false. Such a method is geared toward getting on the attacker’s “level” and demonstrating to her that her grounds for action are invalid. “Invalid,” here, is the key word. As we shall see, at

⁵³ Friedrich Nietzsche, *Thus Spoke Zarathustra*, (Cambridge: Cambridge University Press. 2006).

⁵⁴ Nozick, *Anarchy, State, and Utopia*.

the heart of such a method is the demonstration of formal invalidity and not specifically unsoundness; if we prove that her statements are formally invalid, they are, of course, automatically unsound. To form such a method, we will need some help. Although a host of theorists have proposed various methods of justification, most have done so only within the frameworks of their own respective theories. Aimed, therefore, at demonstrating a positive justificatory framework, many have neglected the importance of establishing, first, a *negative* justificatory framework. The importance of such a framework is this: If we can demonstrate that the claims of nihilists, subjectivists, relativists, and a myriad of other positions are invalid *on their own grounds* and not simply on the specific grounds we develop within our own theories, we can simultaneously demonstrate that it is they who have the burden of proof; they will have to demonstrate both exactly how the grounds *for* their nihilist, subjectivist, and relativist positions are not absurd and how their grounds *against* ethical objectivism are not equally absurd. If we can establish a method of negative justification to class together all of the sorts of statements that are *not* justifications, we can justify the narrowing of our search for a sound ethical model by eliminating such models from consideration. Although there are not many theorists who continue to flirt with these ideas, those most recent theorists who have, including Mackie and Rorty, have contributed significantly enough to them in the modern era that we ought to take up these positions and, through the establishment of a clear method, eliminate them from further consideration.

3.2.2 *Problems in Walton's Model of Justification*

Before we can do this, however, we must contend with some competing paradigms of justification. Although various approaches to ethical justification exist, Douglas Walton's is particularly salient due to the particular depth and rigor with which he has analyzed the subject. In *Ethical Argumentation*, Walton defines ethical justification as "an argument that fulfills a probative function, but not in a conclusive way that guarantees the rational acceptance of the conclusion beyond further doubt or argument."⁵⁵ This definition derives from his overall methodological approach, which is neither deductive nor inductive but abductive, consisting in a chain of layered argumentation governed by what he calls "the probative function." As participants in a dialogue progress down a typical chain of argumentation, the overall goal is to tentatively prove one side or the other, while implicitly acknowledging that the conclusion reached may change upon the discovery of further facts. When deep disagreement occurs, that is when disagreement occurs at the most fundamental level in the chain of argumentation, then only a maieutic process, a process whereby we come merely to better understand the respective positions, can result. The ethical

⁵⁵ Douglas Walton, *Ethical Argumentation*. (Lanham: Lexington Books. 2009): 200.

rules which comprise the major premise, according to him, are either endoxic, in other words based upon our observation of existing norms, or are otherwise emotive in some way via persuasive definitions inherent in all ethical terms. Thus, for Walton, ethical argumentation has both subjective and objective elements. As a result, he concludes that ethical argumentation cannot be inductive or deductive in nature. There are, however, serious problems with Walton's approach. In the first case, he makes little headway toward providing proof that ethical syllogisms are inherently abductive beyond claiming that ethical argumentation is relevantly similar to legal argumentation. Among his main arguments for this position is that, because there are always exceptions to ethical rules and because ethical rules are therefore always defeasible, ethical reasoning must be abductive or, in other words, it must always take the form of an argument to the best explanation based upon a tentative assessment of the facts at hand. It is seriously doubtful, however, whether this argument holds and whether the examples he provides of supposedly abductive ethical reasoning are not merely inductive or deductive when actually applied to ethical argumentation. While the argument that the syllogistic structure he proposes is inductive is admittedly not as promising, it is a strong possibility that that which he proposes, when applied to ethics, is merely a deductive syllogism. The abductive syllogism takes the following form.

- (1) X is Y
- (2) A is an X

(3) Therefore, until it is demonstrated that X is always or is sometimes not Y, it is reasonable to believe that A is Y.

Walton's point is that the major premise "X is Y" seems to be the case but might not be, for there might be exceptions to X's being Y of which we are not yet aware. Applied to ethical argumentation, Walton claims, ethical definitions such as "Theft is the taking of property from someone without his consent" and ethical rules such as "Theft is wrong" take the place of "X is Y." Because these rules always or almost always have exceptions and because such definitions are often disputed, it is obvious, Walton claims, that ethical reasoning is inherently defeasible and thus abductive. Such a conclusion is not, however, as obvious as Walton might like it to seem. It is easy to make the mistake of viewing ethical reasoning or ethical rules as inherently defeasible because ethical reasoning has predicative and attributive elements embedded within its syllogistic structure. As such, these syllogisms are not by any means abductive; the alternating predicative and attributive elements merely make it seem as though ethical principles "have exceptions" and are defeasible. This is not, of course, to state that ethical *rules* themselves may actually be defeasible; it is merely to state that these rules may be defeasible, but their corresponding ethical principles or even a monistic, ultimate principle is probably not defeasible. If such a principle exists, then there is no conflict between the defeasibility of the rules which derive from such a principle, based on their weighting, and the notion that ethical syllogisms are deductive. This also follows if, taking a virtue-theoretical approach, we assume that there is a *summum bonum*. I

shall explain precisely why this is the case. Let us take an ethical syllogism of similar structure to the abductive one above, except let us now change the conclusion to have it become a deductive syllogism. This approach might seem odd, but it will soon prove to be instructive. I propose the following.

- (1) X fulfills Y when X tends toward Y.
- (2) A is X.

(3) A fulfills Y when A tends toward Y.

X, we might note, tends toward (or, in other words, has some probability of yielding) Y if it tends toward Y by an infinitesimal amount. If, therefore, A has a probability of 0.0000001 of yielding Y, then it tends toward Y. If something B has a probability of .5 of yielding Y, then we notice something interesting about the syllogism: B seems also to be X. Now, however, we must ask a seemingly important logical question: “Is it logically possible for one entity, X, to be in two different states at the same time, one at a probability of 0.0000001 of yielding Y and another at a probability of .5 of yielding Y?” We must answer “yes.” This is because, in this example, X is a *class* and has an extension. As it turns out, this syllogism is indeed deductive and indefeasible, but it just is the case that an infinite amount of entities can be classified as X. This kind of syllogism, if applied to a particular context, might have the tendency to confuse, and we can now see why. If we examine the first premise more closely, we can see that, in some sense, it seems defeasible. If we take “tends toward” to imply that one and only one criterion fits this description, then we may easily take some entity B and proudly proclaim, “Aha, the first premise is defeasible because B defeats A.” In this sense, our speaker would be right, but if we are to translate our predicate into its adjectival form, “tendential,” we find that the syllogism may be constructed to refer to either the predicative or attributive usage. As such, we can also see that, interpreted as predicative, A is indeed indefeasible, and so is B, but when we interpret the predicate in an attributive sense, we are forced to do the same thing we are forced to do in ethics: choose one alternative over the other and, yes, defeat some alternative. This, however, does not by any means demonstrate that the syllogism is abductive; it merely shows that value properties are both extrinsic and intrinsic and that intrinsic properties regulate and set the bounds of their corresponding extrinsic properties such that these extrinsic properties can then be defeated within the bounds which the corresponding intrinsic properties have established. Thus, when we want to compare A and B, we also use a deductive method of reasoning, namely the following.

- (1) (a) If $(P(Y|B) > P(Y|A)) \wedge (P(Y|A) \neq 0)$, then both A and B fulfill Y; but (b) If $P(Y|B) > P(Y|A)$, then B fulfills Y to a greater degree.

Lemma: (a) reflects predicative usage of “fulfill,” and (b) reflects attributive usage of “fulfill.”

- (2) $(P(Y|B) > P(Y|A)) \wedge (P(Y|A) \neq 0)$.

(3) (a) A and B fulfill Y; (b) B fulfills Y to a greater degree.

Note here that B defeats A but that both syllogisms I have presented are deductive. Whereas Walton presents an account of ethical discourse in which principles are supposedly defeasible and in which argumentation is abductive, we can see that this approach is neither necessary nor sufficient for ethical argumentation to take place. In fact, I believe that the model which I have provided is the one which actually reflects what occurs in ethical discourse. I believe, as I will show in further chapters, that such argumentation is inherently teleological, as partially suggested by Walton's own chained model of argumentation leading to premises concerning the ultimate, fundamental values of each participant. It is neglect of such possibilities as an ultimate principle or a *summum bonum* which unfortunately mars his account, for I believe that his neglect of the possibility of the existence of such entities leads him to overlook the account of ethical argumentation I have offered. Indeed, such a principle or *summum bonum* would fit neatly into the predicative version of the major premise. Instead of Y, we would speak of "the *summum bonum* Y," or "the ultimate principle Y," and instead of "X," we would use the terms "good," "valuable," "necessary," "obligatory," and such others, and instead of "A" or "B," we would describe actions or traits. Indeed, as I shall prove in later chapters, we already do all of this when we make a value claim.

With regards to emotive terms and persuasive definitions, I have already provided some strong reasons in my analysis of Hare's theory to abandon the idea that ethical terms are inherently emotive. I will grant that it is possible that some words to which we ascribe ethical connotations are emotive, and I will also readily grant that many ethical terms contain emotive meaning, but I cannot agree with the oversimplistic claim that all ethical judgments contain *solely* emotive meaning. Again, we may state that such terms as "murder," "abortion," "courage," and "happiness" evoke emotion, but we must always ask *why* they do so. The answer is simple: they evoke emotion because they are perceived as bad or good, respectively, and this badness and goodness is discerned by a judgment founded upon *reasons*; the goodness or badness of these entities is inherently linked to a reason or explanation. Indeed, we can even use such reasons to change our emotive attitude toward these terms, and we often do. Persuasive definitions can be avoided, however, by speaking in the most general manner about ethics, namely about what one should and must do. One may use all the persuasive definitions one wants, but in the end, the fundamental questions of ethics demolish them: "But why should I value X?," "Why should I think X good or bad?" If these are not fundamental questions of ethics, that is, if "ought" and "good" are loaded terms or persuasive definitions, then I would ask Walton precisely what he means by "ethics."

Persuasive definitions may be a part of ethical discourse, but that is precisely the only claim Walton can make. He cannot, for instance, go from stating "X is how ethical argumentation takes place between participants" to the conclusion "X is how ethical justification should occur." He cannot, in other words, infer how ethical justification should take place from how ethical argumentation between participants normally takes place. This inference, and indeed Walton's entire model, rests upon a series of normative claims concerning how argumentation *should* take place in order to have *good* results, especially with regards to the probative function and the

maieutic process. The only way it seems that he could justify such claims is if he used his own model to do so, but then this would of course be circular and, in Walton's terminology, lack probative effectiveness. My previous claims, therefore, namely that to justify a statement is "to prove it to be true" and that we must begin a program of ethical justification by demonstrating that certain such statements are false, hold.

3.2.3 *The Justificatory Model Explicated and Analyzed*

I will now discuss some of the theoretical models I will be citing in order to develop a negative method of justification, and I will explain their role in my overall model. First, I will examine Hare's theory. Through his justificatory framework, we can demonstrate to the would-be attacker that, if she assumes that preferences are sound bases for ethical judgment, any judgment that she makes to act against *commonly held* preferences, without some further positive justification, is absurd. If our speaker then narrows her judgments to apply to those preferences not shared with her patient so as to attempt a consistent judgment, then we can use our second approach. Using Lillehammer and Hampton's work on internalist reasoning, we can, yet again, demonstrate that, even if one narrows one's judgments in this way, one's judgments are still inconsistent due to the flaws inherent in internalist reasoning. If our speaker then narrows her judgments even further so as not to cite preferences in her maxim at all, but simply her name or some specific attribute she has, we can use Marcus George Singer's adaptation of Kantian theory to demonstrate that such judgments are also inconsistent. The speaker might step back altogether and tell us that she does not acknowledge ethical judgment as such or that she is a nihilist. Although I will demonstrate such a position to be *prima facie* absurd, we will see further, as we delve more deeply into positive dialectical theory, why such a position is absurd. In our exploration of negative justification, therefore, we will see that there are, indeed, certain reasons that cannot be put forth as ethical justifications because of the contradictions, vacuousness, or circularity they entail. It is people who put forth such justifications, we shall see, that have the main burden of proof. Let us start with Hare's theory.

In response to our initial reply to the would-be attacker's reasons for attacking, suppose she responds, "Well, I think I ought to attack and steal from him *because I like to have money*." "Because I like to have money," then, is her given reason. As we shall see, this does not justify her attacking and robbing the would-be victim. Suffice it to say, and for whatever reason, most of us like to have money. *Liking to have money*, then, will probably constitute a *commonly held* preference between the would-be attacker and the would-be victim. Since "Because I like to have money" is her reason and because *liking to have money* is a *commonly held* preference, she will then have to admit, "Anyone who cites 'Because I like to have money' as her reason for attacking and robbing a would-be victim ought to attack and rob him." This is, perhaps, a better way of presenting universal prescriptivism. From "I ought to do X to Y in situation Z,"

it does not clearly follow that “P ought to do X to Y in situation Z.” This is altogether the wrong approach toward universalization. Rather, we must focus on the reason *itself* that is offered for justification since, implicit in the maxim “I ought to do X to Y for reason Z,” it is the reason itself that is posited as a *sufficient condition*. To understand more clearly what I mean, we may present this formally. “I ought to do X to Y for reason (or because of) Z” can be rewritten as “HaveReasonZ(i) \rightarrow OughtToDo(i, x, y).” What this demonstrates is that Reason Z is given as a sufficient reason for the permissibility of her doing X to Y and, also, that “I ought to do X to Y” is a necessary condition of having that reason. Anyone, therefore, who states “I have Reason Z,” on her account, ought to do X to Y since this is posited, once again, as a *sufficient reason*. The “I” and the “me” here are universalizable, since anyone can refer to herself in this way; more specifically, unless she gives additional reasons for her citation of the specific individuals involved, such a citation is arbitrary. This, I believe, is what Hare has tried to highlight, although his non-descriptivist framework has disallowed him from referencing external reasons in the same way that I have. Thus, if she states “I ought to attack and rob him because I like to have money,” she must admit that “Because I like to have money” is a sufficient condition of anyone who likes to have money’s justification of “I ought to attack and rob him.” Thus, the statement “HaveReasonZ(i) \rightarrow OughtToDo(i, x, y),” as she has formulated it, is, indeed, necessarily universalizable unless, of course, she were to designate the “I” to refer *specifically to herself* or the “him” to refer *specifically to the would-be victim*, which she might well do. Of course, she will have to provide additional reasons for the justifiability of such criteria. We will explore this in a bit. First, let us see where universalization of the statement “HaveReasonZ(i) \rightarrow OughtToDo(i, x, y)” takes us. As this statement is necessarily a universalizable one, once again centered on Reason Z as a sufficient condition of the justification of the consequent, she must admit the following: “Anyone who states ‘I ought to attack and rob Y because I like to have money’ ought to attack and rob Y.” Since these are the criteria which she, herself, has set forth, this is what she must accept if she is to be consistent. Since most people share this preference of liking to have money, most people could conceivably utter this statement: “I ought to attack and rob Y because I like to have money.” This time, however, Y could be she. She must admit, then, that if anyone utters this statement with regards to her, that is they replace the variable Y with her, they would be justified in attacking and robbing her. She must, as Hare suggests, be willing to admit the following, “Anyone ought to attack and rob me because she likes to have money.” The grave consequence is this: In attempting to justify an action against someone else’s preferences that happen to be her own preferences, she must covertly acknowledge that she would not mind to have her preferences so acted against. Stated more precisely, she must admit that it is not the case that she prefers to have her preference not-acted-against. Thus, she must implicitly acknowledge, “It is not the case that I do not prefer to be deprived of what I prefer.” Her preference, however, and this goes for all preferences of the same first- or second-order type, is precisely *not* to be deprived of what she prefers; this is inherent in the meaning of “preference.” Herein lies the contradiction inherent in her prescription: She must acknowledge, if she is to be consistent with her original prescription, “It is not the case that I do not prefer to be

deprived of what I do not prefer to be deprived of.” This takes the form of $\neg(P \rightarrow P)$ or, tautologically, $\neg(\neg P \vee P)$. Her prescription, thus, implies a contradiction. Her action of attacking, therefore, entails a pragmatic inconsistency or, if you like, a practical contradiction. Monologically, one could therefore never justify either such an assertion or such a course of action to oneself qua that action as partially embodied by the given reason. If her stated reason is a sufficient condition of someone’s being deprived of this preference, and she has the preference herself, we can confidently say that her reason for acting is unjustifiable. We must be careful, however, to keep in mind what this logically implies. It does not follow from the above argument that any sufficient reasons posited that merely include liking to have money are unjustifiable reasons; it simply follows that liking to have money, by itself, does not constitute a sufficient reason. As we shall see in later sections and chapters, however, liking to have money, as well as reasons relevantly similar to it, is rarely to figure into a justifiable sufficient reason in the first place. For now, let us take on the would-be attacker’s next likely challenge.

To dodge the problems inherent in citing commonly held preferences between her and her patient, she might tell us that she ought to attack and rob the person because he is a cross-dresser. At this point, the reader might point out that she is still acting against some preference that she has and that, in order to avoid an implicitly contradictory judgment based on commonly held preferences, she must still provide some explanation as to how she can justify acting against her own preferences, namely those of not being attacked and robbed. She might, indeed, be stuck by this, but let us, for the sake of argument, assume that she has, somehow, overcome this challenge. We could, of course, simply say that she has decided not to attack and rob him and that she has, instead, decided to call him a slur. This way, we might let her narrowly avoid the former challenge by claiming that she would never and could never have the preference to cross-dress. In any case, let us assume, for the sake of argument, that her attacking the man does not constitute the violation of a commonly held preference. She prescribes, then, “I ought to attack and rob him because he is a cross-dresser.” We might completely undermine this kind of prescription simply by offering the arguments Lillehammer and Hampton have established against internalism (see Sect. 3.1). We do not have to rehash these arguments here. Suffice it to say that, if we undermine the internalist model, we undermine the basis for ethical justification based on statements of desire or preference in general. If the would-be attacker insisted on subscribing to internalist reasoning, we might point out that her reasoning boils down to something along the lines of the following:

- (1) He is a cross-dresser.
- (2) His being a cross-dresser is a sound reason for my attacking and robbing him because it is something I desire or prefer to do.
- (3) My desires or preferences are sufficient reasons for my attacking and robbing him because my desires or preferences are sufficient reasons for my attacking and robbing him.

(4) Therefore, I ought to attack and rob him.

I hope the circularity and vacuousness of this syllogism is apparent enough. Our would-be attacker is now probably getting frustrated. She attempts a last-ditch effort at justification. She states, "I ought to attack and rob him because I am myself." Notice, now, that she is no longer citing, at least directly, any preferences in her prescription. Her only relevant criterion is that she is herself. We might ask her, as we have license, to justify why she is so special that she has such a privilege to attack and rob people. She tries all sorts of justifications: "I am intelligent, and he is stupid," "I am white, and he is black," "I am straight, and he is queer," "I have pink fingernails, and he has blue eyes." We might yet ask her, "And why are these justifying reasons?" Of course, we have perfect license to ask this of her, for, if we set up her premises and conclusion into syllogisms, we would be correctly puzzled as to how such a conclusion could follow, deductively or inductively, from the premises. She might tell us, "I am not the one who has to prove myself here. It is you who have the burden of proof to show that these aren't justifying reasons!" While we actually do not have the burden of proof here, let us take her up on this anyway. In his *Generalization in Ethics*, Marcus George Singer lays out a Neo-Kantian justificatory framework through which, he claims, it can be shown that deontology is the superior ethical model. While I believe he falls far short of this goal, his theory has yielded fruit. His Generalization Principle, modeled after the Categorical Imperative, states that persons may give no justification either (1) that is self-contradictory or (2) through which any other person can equally validly yield different conclusions. Although, within the context of his argument, he states the principle as "If not everyone ought to act or be treated in a certain way, then no one ought to act or be treated in that way without a reason,"⁵⁶ my bifurcation of the principle into two distinctly separate criteria is, as we shall see, highly instructive. Although Singer tailors some parts of his argument toward refuting attempted justifications based on likes or preferences, we can, as we shall shortly see, apply his comments to attempted justification based on personal identification:

The fact that someone does not like people of a certain type would not, by itself, justify him in claiming that they ought to be excluded from some position. The fact that he does not like them would not be a reason, though he might think that it is. For suppose that it were. Then the argument would be that people of type T ought not to be allowed to hold a certain type of position because A does not like people of type T. But this presupposes that anyone A does not like ought not to be allowed to obtain that sort of position.⁵⁷

The self-contradiction, he explains, is apparent. As he states: "Anyone can argue in the same way." He continues:

If there is anyone who does not like the people A does like (himself included), then these people ought to be excluded also. It is possible on these grounds that no one ought to hold any position, and this is absurd. How does A justify his claim to a privileged status for his likes and dislikes? It is possible for him to do so, but not on the ground that he is he, or is a person of 'special importance.' Nor could people of type T justifiably be excluded simply

⁵⁶ Marcus George Singer, *Generalization in Ethics*, (New York: Alfred A. Knopf, Inc. 1961): 31.

⁵⁷ *Ibid.*, 25.

on the ground that they are people of type T. This (or the principle underlying it) also would apply to everybody: people of type U ought to be excluded because they are of type U, and so on.⁵⁸

Singer's comments are of great relevance to our would-be attacker's claims that she is intelligent, white, and so on, and, thus, that she ought to attack and rob her would-be victim. "Fine," she admits, "I won't give you the satisfaction. I won't give you any additional reasons for why my being me privileges me to attacking and robbing him. All I will insist on is that *I am me*, and that is all there is to it; my being me is all the justification I need." But is it? The same point Singer makes above can be applied to this claim, for, to repeat Singer's wise phrase, "Anyone can argue in the same way." I can argue that I am me and come to the opposite conclusion, namely that I ought not to attack and rob the would-be victim. "Wait a moment," she tells us. "That violates the law of identity; I can never be you, and you can never be me." She might have a point, but she is still flat wrong. How so? I can argue that she is herself and still come to the opposite conclusion, and I can do so with equal validity. This constitutes a counterexample; her argument is invalid. Let us see this in a syllogism:

1. I, Ms. X, am myself.

2. Therefore, I ought to attack and rob Y.

Counterexample:

1. She, Ms. X, is herself.

2. Therefore, she ought not attack and rob Y.

Thus, we have shown several types of common attempts at ethical justification which plainly fail on grounds of formal invalidity. This is important, for, although a very simple argument, it corroborates Michael Smith's argument that value judgment is a species of statement of belief, rather than a statement of desire.⁵⁹ Thus, on his account, although we might have a desiderative motivating reason to do something, this motivating reason does not and could never constitute a normative reason to do that thing. My demonstration of negative justification has been important precisely because it provides further proof for Smith's claim concerning value as belief; as I have shown, in order for one to claim that one ought to do anything at all, even prudentially, one must stand up to the test of the syllogism, the validity or soundness of which we must either believe or disbelieve. As a consequence, my demonstration of negative justification corroborates many other philosophers who have made similar moral-psychological claims concerning value and belief.⁶⁰

⁵⁸ Ibid.

⁵⁹ Smith, *The Moral Problem*.

⁶⁰ Foot, "Moral Beliefs"; *Natural Goodness*; Hampton, *The Authority of Reason*; Rosalind Hursthouse, *On Virtue Ethics*, (Oxford: Oxford University Press. 1999); John McDowell, *Mind, Value, & Reality*, (Cambridge: Harvard University Press. 2001).

Interestingly, Derek Parfit's approach to ethical justification is also somewhat similar to what I have espoused, namely in his use of logical counterexample.⁶¹ In his case, it is the features of Critical Present-Aim Theory and morality that he uses in tandem to demonstrate the weaknesses of Self-Interest Theory. Through a complex analysis of personal identity, he demonstrates that Self-Interest Theory must be unsound on its own terms. Leaving aside the soundness or unsoundness of his theory of personal identity, Parfit successfully shows that Self-Interest Theory must be unsound because, in order to avoid charges of inconsistency, Self-Interest Theorists must maintain temporal neutrality; that is, they must consider equally weighty both their present and futures interests. This, he explains, is because Present-Aim Theory and Critical Present-Aim Theory, respectively, present reasons to agents which are equally as valid (also meaning potentially as invalid) as the reasons for action which Self-Interest Theory presents to agents. In the case of Present-Aim Theory, the reason in question is the agent's present desire, and in the case of Critical Present-Aim Theory, the reason in question is the overriding reason bearing on the agent to act which might coincide with or which might be contrary to the agent's present desires. Since these theories present to the agent reasons for action which are equally as valid as those set forth in Self-Interest Theory, the Self-Interest Theory is exposed as implausible. If the Self-Interest Theorist claims that pursuing solely her or his interests is supremely rational, then she or he must also accept as supremely rational the Critical Present-Aim theorist's potential claim that she or he has an overriding, supremely rational reason to act against her or his own interests. This, Parfit explains, is because, if the Self-Interest theorist attempts to refute this claim, she or he can rightly be accused of inconsistency: the Critical Present-Aim theorist can claim, rightly, that the Self-Interest theorist is being inconsistent in claiming that she or he is being supremely rational by taking a strongly particularist position on *her own* interests while considering irrational the Critical Present-Aim Theorists' and the Present-Aim Theorists' strongly particularist position on his *present* overriding reasons or desires, respectively. Thus, in order for the Self-Interest Theorist's position to be consistent, he must accept temporal neutrality or the supreme rationality of giving equal weight to present and future interests. If the Self-Interest Theorist accepts this, however, the Self-Interest Theory fails, for it fails to make any strong claims to the supremacy of the reasons it gives for action. As one might again note, this is relevantly similar to the syllogistic approach I have taken above.

Our attacker might insist, however, that we must prove why she cannot simply be a nihilist and claim that she does not have to make value judgments at all or that there are actually such things as justifiable ethical judgments. She might point out, and she would seem to be right, that we have thus far only demonstrated that there are such things as unjustifiable ethical judgments. That, however, is simply the point: nihilism's essential proposition is that ethical statements cannot have truth

⁶¹ Derek Parfit, *Reasons and Persons*, (Oxford: Oxford University Press. 1986); *On What Matters: Volume I; On What Matters: Volume II*, (Oxford: Oxford University Press. 2011).

values.⁶² We have, however, seen this to be false by the simple fact that there are such things as false ethical judgments, that is, false due to the invalidity of the syllogisms from which they derive. Let us be clear, however, concerning what exact sort of value statement I am claiming can be shown to be false. Clearly, I cannot mean to imply that, if a syllogism is subject to counterexample, its conclusion is false; this would be a patent logical fallacy. The sort of statement I am claiming is false, rather, is a statement of the sort “*X* is a valid reason for doing *Y*,” where such a reason has been shown, as in the above examples, to be invalid. Reasons, we must recall, are fundamentally normative and thus invoke value and necessitate value judgment; the citation of a reason is essentially the citation of an “ought” or, alternatively, the citation of a better course of action. There indeed exist, therefore, false value judgments.⁶³ In the next two sections, we will further counter the challenge of nihilism and demonstrate how Gewirthian rationalism and Habermasian deontology stand up to the task of *positive* justification. In later chapters, we will also demonstrate how virtue theory underlies such a body of justification. For now, however, let us make sense of this section’s analyses. Our goal, as one might recall, was to establish a method of negative justification. As a result of the previous analyses, we are now in the position to lay out some general rules for justification. They are as follows:

Rule 1—Anti-Internalism: The premises upon which prescriptions are based may not contain any antecedent or, otherwise, sufficient conditions that make reference to preference or desire alone. (Lillehammer, Hampton)

Corollary: *If an agent’s sufficient reason does contain a reference to preferences or desires, it is still invalid if it violates the commonly-held-preferences interpretation of Harean universal prescriptivism. (Hare)*

Example: “I desire or prefer *X*, therefore I ought to do *Y*.”

Rule 2—Validity: The prescriptions which an agent proposes may not be based on syllogisms whose premises or conclusion are subject to counterexamples. (Singer, Kant, Parfit)

Example: “I am me/I am race *X*/I am gender *X*, therefore I ought to do *Y*.”

Rule 3—Non-Self-Contradiction: The prescriptions which one proposes must not be self-contradictory; that is, they may not contradict themselves, neither explicitly nor implicitly. (Singer, Korsgaard, Kant)

⁶² This, of course, is one use of the term “nihilism.” Another interpretation of nihilism holds that ethical statements have truth values but that they are all false. This, however, is not the interpretation I am addressing at present.

⁶³ Of course, this also implies that there can be true value judgments of the form “*X* is not a valid reason for doing *Y*,” but this is a somewhat vapid point, for the statement here is still essentially negative in that it merely further clarifies which reasons are not valid and not, by contrast, which reasons are valid. A positive theory of value judgment would demonstrate not only which kinds of value statements are true but, in addition, which kinds of statements yield valid reasons for action.

Example: “I do not want people of type X to work in position Y, therefore I ought not to hire people of type X.”

Rule 4—Non-Circularity: One’s prescription may not be based upon premises which presuppose the prescription, and the premises one uses to form one’s prescription may not presuppose themselves.

Example: “X is a valid basis upon which to prescribe Y because X is a valid basis upon which to prescribe Y. Therefore, I ought to do Y.”

Rule 5—Non-Regression: One’s prescription may not found itself upon infinite regression. (Walton)

Example: “All values are created by X, and X is a value.”

All of the arguments to which I have objected violate one or more of these rules. With these rules and their analytical bases in hand, let us see what sort of method they yield. The method is centered on providing a solid rubric that will help us move step by step toward evaluating and eliminating proposed bases for justification and which will, ultimately, push us in the direction of *positive* justification. The rubric is as follows:

I. Establishing the Main Burden of Proof

- *You have the main burden of proof if*
 1. You cannot demonstrate your positive argument to be valid or sound (inclusive disjunction), but your opponent can demonstrate it to be invalid or unsound (inclusive disjunction).
 2. You can demonstrate your positive argument to be valid, but your opponent can demonstrate it to be unsound.
- *You do not have the main burden of proof if*
 1. You can demonstrate your positive argument to be valid or sound (inclusive disjunction), and your opponent cannot demonstrate his positive argument to be valid or sound (inclusive disjunction).
 2. You cannot demonstrate your positive argument to be valid or sound (inclusive disjunction), but you can demonstrate your opponent’s positive argument to be invalid or unsound (inclusive disjunction).
- *Person who Fulfills Criteria for Having the Main Burden of Proof: The Attacker*
 - *Why?:* She has violated rules 1–5 via internalist reasoning, etc., and we have not violated any rules.
- *Person who Fulfills Criteria for Not Having the Main Burden of Proof: Us*
 - *Why?:* We have not violated any rules, and the would-be attacker has violated rules 1–5.

II. The Charge of Vacuousness

– *Your account is vacuous if*

1. You do not have the main burden of proof, but you have not given a valid or sound *positive* account of ethical justification.
2. You do have the main burden of proof, and you have not given a valid or sound *positive* account of ethical justification.

Corollary: *By analytic consequence, you cannot simultaneously have the main burden of proof and have provided a sound positive account of ethical justification.*

– *Your account is not vacuous if*

1. You do not have the main burden of proof, and you have given a valid or sound *positive* account of ethical justification.

– *Person who Fulfills Criteria for Having a Vacuous Account:* Both the Attacker and Us.

At this point, the reader might be wondering, “How is it that we do not have the main burden of proof, our opponent does have the main burden of proof, but both of us have vacuous accounts?” This is because the main burden of proof concerns which person is able to demonstrate a *negative* account of his opponent’s *positive* account. It does not, however, concern who is actually able to provide a *sound positive account*. Since neither of us has established a sound positive account of ethical justification, one of us must proceed toward one if at least one of our accounts is to prove non-vacuous and if we are to prove that sound positive ethical justification is possible. In order to do this, we must first evaluate the two positions according to the criteria in the above-established rubric and rules. Of the two accounts, we must judge for which one or ones it is logically possible for there to exist a sound positive account or, at least, for which one or ones it is not logically impossible for there to exist a sound positive account. Since, as I pointed out above, “You cannot simultaneously have the main burden of proof and have provided a *sound positive account* of ethical justification,” it should be obvious that it is logically impossible for the attacker to develop, on her current grounds, a sound positive account of ethical justification.

Let us now dig a bit deeper. What proof can we offer to our attacker, as well as nihilists or anti-realists in general, that there exists a sound positive account of ethical justification? How can we prove that such judgments as “I ought to do X” and “X is good” have truth values that can be objectively evaluated? It is precisely here that the concept of rationality returns to the foreground. In the next section, we will consider the major merits of Gewirth’s theory. In the section following this, we will then consider the theoretical insights of Habermas’s communicative ethics. As we shall see, it these two accounts that provide us with a mutually supportive and highly substantive method of evaluating the truth value of such statements as “I ought to do X.” As we shall see to some extent in Chap. 4 and, to a larger extent in Chap. 5, it is virtue theory that provides the most substantial model for evaluating the truth value of such statements as “X is good.” Contrary to Gewirth’s view that all sound virtue

ethics must be based on rationalistic deontology, we shall see that the true state of affairs is the reverse. In the end, we shall see that dialectical deontology, while highly powerful in the domains in which it is operative, is wholly dependent upon the soundness of the premises of a virtue-theoretical model of value.

3.3 A Sound Positive Account, Part I: An Analysis of Gewirth's Ethical Rationalism

In this section, I analyze Gewirth's dialectical-deontological rationalist theory. I corroborate the claims of such other theorists as Deryck Beyleveld and Roger Brownsword in demonstrating the overall soundness of Gewirth's theory. While some of Gewirth's supporters, especially Beyleveld, have explored in depth the consequences of Gewirth's theory for applied ethics and various moral dilemmas, I employ an altogether different method of analysis of Gewirth's theory in order to assess such applications, one that I believe most clearly and directly derives from his theory. As such, I believe it demonstrates most clearly what conclusions pertaining to applied ethics are derived from his theory; as such, the conclusions I reach concerning the practical applications of Gewirthian theory are somewhat different. Like Beyleveld, I seek partially to demonstrate the power of Gewirth's theory by elucidating its applications, but I seek to extend Beyleveld's own analyses. As such, I demonstrate that Gewirth's theory has extremely wide application and that it stands up to dilemmas of all kinds, as well as to analyses of animal rights and ethics. While sound and deeply substantive, the theory has a fundamental flaw: Gewirth's use of the terms "ought," "must," and "good" lack substantive analysis. As such, his theory fails to tell us anything about good states of affairs or good things more generally. Lacking teleology, his theory proves to be one-dimensional, disregarding of anything but the relevance of human action, and thus vapid in certain ways. His theory is, thus, formally substantive but in many ways contentually weak. As such, his account suffers from what Slote calls "self-other asymmetry." I go even further with this critique in offering what I call the Single-Person Problem, which I claim severely undermines all non-virtue-based theories. The insights from Gewirth's dialectical method, however, are absolutely vital for their elucidation of the structure of value judgment. Connecting this elucidation to my analysis and characterization of rationality, I conclude that the linguistic-conceptual framework that rationality constitutes is further revealed through Gewirth's theory. Taking these insights, I move to an analysis of Habermas.

3.3.1 The Basic Elements of Gewirthian Theory

Much work has been undertaken on Gewirth's theory since his publication of *Reason and Morality* in 1978. Numerous journal articles, an anthology, and even a

comprehensive step-by-step defense have entered into the philosophical fray. Beyleveld's *The Dialectical Necessity of Morality* is possibly even clearer and more comprehensive than Gewirth's original work, due to its in-depth analysis of each premise and his comprehensive replies to objections at each step. As a result, an in-depth analysis on my part would be superfluous and repetitive. Instead, I will simply lay out the premises of the theory for the reader and provide some minor correction to certain word usage and other minor details in the theory with which I disagree. I urge the interested reader to explore these works and the theory itself, as well as objections and replies to the objections.⁶⁴ Since Beyleveld and others have dealt with possibly every objection that has ever been offered or ever could be offered, I will eschew objections and replies here; they can be found in these works. My aim in this section is to demonstrate that, not only is Gewirth's theory a sound positive account, but also that it is a good theory in that it is widely applicable. Meta-ethical theories which seem sound on a highly theoretical level sometimes break down in practice. They sometimes fail when it comes to answering such concrete questions as "How should we treat non-human animals?," "Is voluntary active euthanasia permissible?," or "What should I do in moral dilemma X?" Gewirth's theory, I will show, deals adequately with all such questions and, as such, does not break down in practice.

Gewirth's theory, like Hare's, is dialectical. As such, its premises are ones formed by statements made from the perspective of the agent. Unlike Hare's theory, however, Gewirth's is dialectically *necessary*. This means that the theory is not constituted by statements that agents happen to make or incidentally make; rather, it is constituted by statements that agents must make simply by virtue of their being agents. In other words, Gewirth claims that there are certain statements that all agents make simply because they are agents. Although at first seemingly inconsequential and glaringly obvious, Gewirth highlights an important point: In order to engage in action of any sort, agents must *act* and *judge*. He therefore puts agency itself in the ethical spotlight; agency and rationality, he suggests, are the single-most relevant criteria to ethics.

Gewirth's theory may also be considered, to some extent, an offshoot of discourse theory. Although Habermas rejects theories which take "the form of a hypothetical process of argumentation occurring in the individual mind"⁶⁵ in favor of those which are derived from actual, practical discourse, Habermas himself emphasizes that we can only understand the ethical significance of practical

⁶⁴ Gewirth, *Reason and Morality*; *Human Rights: Essays on Justification and Applications*, (Chicago: University of Chicago Press. 1983); *The Community of Rights; Self-Fulfillment*, (Princeton: Princeton University Press. 1998); Regis, *Gewirth's Ethical Rationalism*; Beyleveld, *The Dialectical Necessity of Morality*.

⁶⁵ Habermas, *MCCA*, 68.

discourse if we analyze its necessary preconditions.⁶⁶ Cannot this only be achieved by observation of others' actions and development? Unlike Habermas, Gewirth implicitly offers a foundational grounding to discourse theory, and this, we should acknowledge, is necessary. In order to establish and fully understand the necessary preconditions of moral discourse and communicative action, must we not establish an accurate account of action more generally? In order to answer such questions as "What is right action?" and "Why ought I act rightly?," we must first answer the underlying question "What is action?" While this question is inherently discursive and dialogical, Gewirth, unlike Habermas, points out that certain conceptual components underlying this question also exist which are indeed fundamentally monological.

Gewirth uses the specific term "action" to refer to the generic processes and results of agency. While he does clearly distinguish "action" from involuntary actions, impulses, and compulsions, I believe a more accurate term to describe the generic processes and results of agency is "voluntary action," which, in any case, is a term he uses where he deems it necessary. While this expression "generic processes and results of agency," which is in any case *not* an expression which Gewirth uses, is a bit vague, we might recollect the definition of rationality that I have set forth, that is

$$\text{HasLogicalRationality}(\Phi) \leftrightarrow \text{HasExperientialRationality}(\Phi)$$

or rewritten:

$$\begin{aligned} &\text{HasLogicalJudgment}(\Phi) \\ \leftrightarrow & (\text{HasConceptualSynthesis}(\Phi) \wedge \text{HasConceptualAbstraction}(\Phi) \wedge \text{HasFreedom}(\Phi)) \end{aligned}$$

⁶⁶ Ibid.

86: "Every person who accepts the universal and necessary communicative presuppositions of argumentative speech and who knows what it means to justify a norm of action implicitly presupposes as valid the principle of universalization. . ."

89: "If these considerations are to amount to more than a definition favoring an ideal form of communication and thus prejudging everything else, we must show that these rules of discourse are not mere *conventions*; rather, they are inescapable presuppositions."

See also his application of Kohlberg's theory of moral development:

170: "These clarifications proved helpful in the attempt to ground moral stages in a logic of development. Kohlberg's social perspectives are intended to have this function. As we saw, they can be correlated with stages of interaction that are ordered hierarchically according to perspective structures and basic concepts. This allows us to see how notions of justice are derived from the forms of reciprocity available at the various stages of interaction. With the transition from normatively regulated action to practical discourse, the basic concepts of principled morality spring directly from the reorganization of the available sociocognitive inventory, a reorganization that occurs with the necessity of developmental logic. This step marks the moralization of the social world, with forms of reciprocity that are built into social interaction and become increasingly abstract forming the naturalistic core, so to speak, of moral consciousness."

I urge the reader to review these constituent concepts as they have been presented in Sect. 2.3. As rationality is necessary and sufficient for agency, when we speak of rationality and agency, we should expect a lot of overlap. Such overlap is apparent in Gewirth's account of what he calls the "normative structure" of agency. He states:

Because of its generic features, action has what I shall call a 'normative structure,' in that evaluative and deontic judgments on the part of agents are logically implicit in all action; and when these judgments are subjected to certain rational requirements, a certain normative moral principle logically follows from them. To put it otherwise: Any agent, simply by virtue of being an agent, must admit, on pain of self-contradiction, that he ought to act in certain determinate ways.⁶⁷

Gewirth's account of the generic features of agency have much in common with the definition of rationality that I have given. In fact, we might go so far as to state that his characterization of agency is a characterization of a particular type or conception of rationality, namely practical rationality. We may sum up Gewirth's characterization of the generic features of agency as follows:

1. Voluntariness or freedom
2. Value judgment
3. Purposiveness or Intentionality

As I have explained in my own account of rationality, voluntariness or freedom is a key component, but the manner in which Gewirth attempts to elucidate the nature of freedom is in a merely judgment-based manner, which, as I have stated, leads to the neglect of the questions and problems I have highlighted thus far throughout this work. As we shall see, however, the elucidation of merely this aspect of freedom is instructive and highly powerful. In the first case, it is tautological to state that all of our voluntary actions, that is those which are relevant to ethics, are done voluntarily. That is clear enough. Deliberation and choice, however, are characteristics that are a crucial part of such voluntary action. Deliberation, on his account, is merely the concentration toward a series of possible choices. Via deliberation, choices are those things which are "picked out" for action. It is within these phenomena of deliberation and choice, in other words within the phenomenon of logical judgment, that Gewirth thus finds some of the simplest yet most startlingly relevant and important claims made on the part of agents.

Gewirth's account of value judgment is, perhaps, the single-most important feature of his theory. According to Gewirth, there is no such thing as a voluntary action without value judgment. As he states: "Every agent implicitly makes value judgments—there are no indifferent rational actions."⁶⁸ Considering my own characterization of rationality, it seems clear enough that there must be at least *some* judgment involved in voluntary action. You might try to reflect more deeply upon this. Try to act without making some value judgment: tie your shoes, walk

⁶⁷ Gewirth, *Reason and Morality*, 26.

⁶⁸ *Ibid.*, 40.

across the room, fling a piece of paper across a table, deposit money into your bank account, go to bed. If done voluntarily, these are all the result of value judgments, no matter how insignificant, implicit, or even subconscious.

Purposiveness or intentionality is also central to his theory. All voluntary action, he states, is purpose- or end-directed. Even if the agent in question acts in manner X only for the purpose of acting in manner X, this is still purposeful behavior, although the act and the purpose might be the same. Thus, in an important sense, all action is teleological; it is, as Aristotle stated, always aimed at some good, or to be less controversial and question-begging at this step, at some *thing*. It is this aspect of Gewirth's theory that I find particularly important. Although, in Gewirth's theory, the teleological nature of voluntary action has consequences only for elucidating the structure of value judgment and thereby helping to constitute one of Gewirth's main premises concerning such judgment, it is again in failing to substantively analyze the consequences of intentionality experientially that Gewirth's theory falls short of applying to the questions and problems I have mentioned. Had Gewirth undertaken such an analysis, as I will later, he might have realized that the concept of a *telos* of action has much greater and more far-reaching consequences for the concept of value.

We might consider all of the counterarguments that have been hurled against this characterization. A few examples include: "Free will is an illusion," "One need not deliberate in order to make a choice," "Agents need not make value judgments to act voluntarily," "Agency has been different in different time periods and cultures," "Voluntary action need not be executed for a purpose or, otherwise, be teleological." Again, my refutation of these objections would be superfluous; there already exists a great body of work to refute these. Needless to say, all of these objections, as well as others which have more recently been presented but which rest on premises similar to those already refuted, fail.⁶⁹ As we can even see from my characterization of rationality, having cited various pieces of psychological and neurophysiological evidence, voluntary action does, indeed, have these generic features.

⁶⁹ Deryck Beyleveld, "A Reply to Marcus G. Singer on Gewirth, Beyleveld and Dialectical Necessity," *Ratio Juris* 15, no. 4 (2002): 458–73; Marcus George Singer, "Gewirth, Beyleveld, and Dialectical Necessity," *Ratio Juris*. 13, no. 2 (2000): 177–195. Singer's objections in his "Gewirth, Beyleveld, and Dialectical Necessity" seem especially biting, but they are deeply wrong-headed. He points out a problem in Beyleveld's presentation of Gewirth's theory which I, as well as Beyleveld, acknowledge as problematic. The inference $((P \wedge Q) \rightarrow R) \rightarrow (P \rightarrow (Q \rightarrow R))$, interpreted as a conclusion of logical entailment from material implication, would, indeed, be a fallacy of transposition. This inference, however, is taken out of context and is not representative of Beyleveld or Gewirth's arguments. Thus, Singer's conclusion that there is a "fallacy lying at the root of Gewirth's argument," admittedly predicated on the assumption that "Gewirth's argument has been symbolized correctly [by Beyleveld]," is unwarranted. In his reply (Beyleveld 2002), Beyleveld points out that, if P and Q entail each other, which they do in the Argument to the Sufficiency of Agency, then the above formulation is merely the principle of *modus ponendo ponens*, a necessarily valid principle of the form $((P \rightarrow R) \rightarrow (P \rightarrow (P \rightarrow R)))$. The distinction between material implication and logical entailment here is, thus, moot.

Let us continue by noticing something interesting. Gewirth's account is a kind of naturalist account; the goal of his theory is to derive *values* outlined by what he calls the Principle of Generic Consistency from the *fact* of agency. Inherent in his premises is a statement of fact, and inherent in his conclusion is a statement of value. What, then, follows logically from this fact of agency?

3.3.2 *The Premises and Conclusion Reconstructed and Analyzed*

Let us first, as Gewirth does, consider intentionality or purposiveness. This phenomenon is linguistically represented as "I do X for purpose E." All agents, he claims, at least implicitly assert this when they choose to act. X, as previously stated, may be identical to E. Next, let us consider value judgment. The basic language of value judgment includes "ought," "should," "must" and especially "good." If I choose to do X for purpose E, I have implicitly made a value judgment, namely "I ought to do X," "I should do X," or "I must do X." Since the function of deliberation is to pick out or choose some course of action X, she who deliberates must make a value judgment concerning what she ought to do and, by some judgment of value, eventually judge that she ought to do X. How does she pick out X, specifically, as that which she ought to do? She makes a judgment of goodness. "X," she judges, "is better than some alternative Y." She might also judge "X is less worse than some alternative Y." In other words, the agent is judging her action, in one sense or another, to be good; the judgment might, for instance, be made on the basis of some first-order desire. Even in situations in which someone voluntarily acts self-destructively, she must still implicitly judge that self-destructiveness to be a better choice than some alternative; although she might judge her self-destructiveness to be bad, she must *value* it more than she values some other alternative and, thus, she must at least judge her self-destructiveness to be less worse, less disvaluable, or, perhaps, more important, than not to be self-destructive. Even if someone acts self-destructively, she must still, above all, value her purposiveness; she must judge her purposiveness to be good, for, even if her purposes are to be self-destructive, she must value the purposiveness which enables her to act so. What this means is *not* that an agent cannot make the claim that she wants to do something she thinks bad for herself or morally bad, for people can clearly advocate actions that they believe to be so. What Gewirth and Beyleveld have tried to communicate, and what I believe various critics of this position have overlooked, is the following idea: that all agents who act voluntarily choose, by definition, and, in order to choose, necessarily make choices from *some* motivation, be it desire, conscious value judgment, or mere apathy. When an agent makes the choice to act in a certain way, he employs *some* reason, either as the source or the result of the given motivation. When an agent chooses to act, it is the *reasons* for the determination of these choices for which the agent necessarily stakes a normative claim, that, in other words, the agent necessarily views as prudentially or morally good, better, less worse, and the like. Thus, while it is *conceivable* that an agent might

state “I do not value my purposes, freedom, or well-being,” it is, upon reflection, inherently inconsistent for such an agent to make this claim, for in *making the choice* to make this statement and to participate in such dialectic, and in *concocting a reason* for making the statement and participating in such dialectic, even if the reason is simply “because I want to,” the agent necessarily makes a judgment concerning the value of that reason, even if it is simply a judgment concerning the prudential value of that reason. Thus, while it is conceivable for an agent to *claim* that he disvalues his freedom, well-being, or purposes, it is inconceivable for an agent to *actually disvalue* these things. I believe that Gewirth and Beyleveld have not clearly communicated this in the past; such philosophers as Marcus George Singer and E.J. Bond have thus criticized this Gewirthian view on a set of false presumptions concerning what Gewirth and his supporters are actually claiming.⁷⁰ In fact, if we closely consider Gewirth’s argument concerning the link between motivation and the inherent normativity of reasons, we see that it is actually very similar to Nagel’s.⁷¹ This takes us to our next step in Gewirth’s argument.

Since we have established that all agents must state “I do X for purpose E,” “I ought to do X,” and “X is good,” we can now approach the concept of voluntariness or freedom inherent in agency. Since the latter two statements require voluntariness or freedom, they entail, respectively, “I ought to be free to do X” and “My freedom to do X is good.” In order to act at all, whatever one’s purposes, one must view one’s freedom as a necessarily good means to that purpose or end. These statements, then, entail the judgment, “My freedom is a necessary good.” Gewirth, specifically, tells us that all agents must consider their freedom *and* well-being to be necessary goods, since basic well-being (life, food, etc.) is also necessary for the pursuit of any purposes whatsoever. I, however, will condense these all under the label “freedom,” since these goods are inextricably bound up with one’s freedom anyway. One may cite both freedom *and* well-being, but I do not think this separation is necessary. If we are citing things that are necessary for any action or purpose-fulfillment whatsoever, which we are, we may acceptably class Gewirth’s definition of well-being within his definition of freedom. This idea of well-being as freedom is not new; Amartya Sen has comprehensively explored and corroborated this in opposition to mainstream game-theoretical and rational-decision-theoretical conceptions of well-being and self-interest.⁷² Remember, once again, that all of the analysis we are doing of agency is *not* assertoric; that is, when he states that all agents must judge that their freedom is a necessary good, Gewirth is not making the claim “Freedom is a necessary good.” All he is saying is that all agents *must* make this judgment in order to be agents.

⁷⁰ Edward J. Bond, “Gewirth on Reason and Morality,” *Metaphilosophy* 11 (1980): 36–53; “Reply to Gewirth,” *Metaphilosophy* 11 (1980): 70–5; Singer, “Gewirth, Beyleveld, and Dialectical Necessity,” *Ratio Juris*. 13, no. 2 (2000): 177–195.

⁷¹ Thomas Nagel, *The Possibility of Altruism*, (Princeton: Princeton University Press. 1970).

⁷² Sen, *Rationality and Freedom; The Idea of Justice*.

Although Sen rejects the idea of liberty as mere control (what he labels “direct liberty”), he strongly identifies liberty with the idea of “capabilities.”

Let us move on to Gewirth's next premises. Since all agents must judge their freedom and basic well-being to be necessary goods, that is, goods without which any other good or goods would be impossible for them, they must all also make the judgment "I must have freedom and well-being." This judgment then entails "My freedom and basic well-being must not be interfered with or impeded," which, in turn, entails "Other agents must at least not interfere with or impede my having freedom and well-being." In any purpose agents choose to undertake and which agents judge they ought to undertake, therefore, they implicitly make a demand on other agents not to interfere with their fulfillment of that purpose. If agents did not make this demand, then they would have to accept the judgment "It is not the case that I ought to do X for purpose E." If an agent has chosen to pursue a purpose E via means X, this means, once again, that agents must judge that they ought to do X for purpose E. If they did not make a demand on others not to interfere with their doing X for purpose E, at least implicitly, then they would simultaneously implicitly assert "I ought to do X for purpose E" and "It is not the case that I ought to do X for purpose E." As a result, they would be asserting that they are not an agent, since we have established that all agents must assert "I ought to do X for purpose E." The conditional $(\text{IsAgent}(\Phi) \rightarrow \text{AssertsIOughtToDoX}(\Phi))$ demonstrates this; the negation of the consequent results in the obligatory acceptance of the negated antecedent:

$$(((P \rightarrow Q) \wedge \neg Q) \rightarrow (((\neg P \vee Q) \wedge \neg Q) \rightarrow \neg P)).$$

Thus, all agents do make such a demand. To assert that one does not make such a demand would be an assertion that one is not an agent, which is an obvious contradiction.

Thus, we have established that agents, simply by acting voluntarily and purposively, impose demands on other agents. In other words, they assert that agents *must* refrain from doing something; this is tantamount to making a duty claim. All agents, therefore, make duty claims. Since duties and rights are correlative, that is, that an assertion of a duty entails the assertion of a right and vice versa, agents also simultaneously make a right claim. If agents' duty claim directed toward other agents takes the form "All agents have a duty to refrain from depriving me of X," then the right claim which follows from this is "I have a right against all agents not to be deprived of X." I realize I am proceeding quite quickly but, again, a more in-depth analysis of each step is superfluous, since it has already been undertaken *in extensio* elsewhere. I am simply laying out the general framework of the theory and am trying to elucidate points which I believe the reader might need elucidated. Once again, any objections and replies to these objections are available in the aforementioned sources.

At this point, then, we have established that all agents must make duty claims and correlative right claims, at least implicitly if not explicitly. Our next step is universalization. Let us first elucidate the form of the right claim. From what we have thus far demonstrated, it should be evident that agent right claims take the

following form: “I have a right to X by virtue of Y,” where X is “freedom to pursue my chosen purposes” and Y is “my being an agent.” The criterion “my being an agent” is the criterion for such right claims because such right claims derive from one’s being an agent. Therefore, any given agent must assert “I have a right to freedom to pursue my chosen purposes by virtue of my being an agent.” This agent then, perhaps, realizes that *all* agents, and not simply she, make this assertion; the sufficient reason for having an agent right is being an agent. If, then, she accepts that all other agents make the same right claims as she, yet she does not accept these right claims and, worse yet, chooses to violate other agents’ asserted right to freedom, which encompasses a myriad of different right claims, she implicitly makes the assertion that there exists an agent that does not have a right to freedom since she asserts “It is not the case that some agent ought to be free to pursue some purpose E.” But how can this be? First, she accepts that she has a right to freedom due to her being an agent. However, by interfering with the other agent’s freedom to pursue some purpose E, she must accept that there exists some agent that does not have such rights due to its being an agent; in other words, she accepts “It is not the case that, if something is an agent, then it has a right to freedom.” If she accepts this assertion, then she must likewise accept that she does not have a right to freedom if she is an agent, which she is. This, as before, implies her assertion of a contradiction, since she negates the necessary condition of being an agent, namely claiming a right to freedom. If she is to accept that she is an agent, she must, then, accept that she has a duty not to interfere with agents’ rights to freedom. From these premises, Gewirth derives the following, monistic, ultimate principle, which he terms The Principle of Generic Consistency: *Act in accord with the generic rights of your recipients as well as of yourself.* Once again, several objections could arise, including but not limited to: “Rights are a modern/Enlightenment/Western/anachronistic social construction,” “No rights can be derived from mere demands,” “Agents can disvalue their freedom,” “Agents can make right claims about themselves but need not accept those of other agents,” “Ethical norms cannot be derived from facts.” These, once again, have been thoroughly addressed and refuted by Beyleveld and others. The last objection, especially, should be by now obviously false.

Thus, if I am an agent, which I am, I must not only assert that I have a right to freedom but that I also have a duty not to violate other agents’ rights to freedom. Likewise, all other agents have the same rights and duties simply by virtue of their being agents. Some important questions might now arise: “What about animal rights?,” “What about positive agent rights?,” and “What about moral dilemmas and self-defense?” We will discuss all of these questions in some depth.

3.3.3 Applications to Animal Ethics and the Principle of Proportionality

There are at least three ways in which we might derive animal rights from dialectical theory, all three of them promising and indeed coincident in certain

respects. The first method I will examine is one derived from the premises of Harean universal prescriptivism as I have reconstructed them. I will then explore the model Beyleveld has derived from Gewirthian theory and will suggest that these conclusions, as they are derived from the PGC, yet again neglect highly relevant aspects of virtue theory, partially due to their neglect of the theoretical consequences of a serious consideration of experiential rationality. A serious consideration of these elements, as I will show, suggests, contrary to Beyleveld's admittedly correct derivation from Gewirthian theory, the inherent value of non-agents. The third method I will suggest, therefore, will be a virtue-theoretical one. It is through such an analysis that I hope to further demonstrate the inadequacies of the dialectical approach in the absence of a virtue-theoretical foundation. Of course, this mere suggestion by itself is insubstantial; in order to clearly demonstrate why such an approach succeeds, we will have to wait until we arrive at the more advanced stages of my overall argument. For now, a formal (in the sense of being mostly non-contentual) argument will have to suffice.

The first method by which we could conceivably derive animal rights is via the use of the commonly-held-preferences interpretation of Harean universal prescriptivism. We previously demonstrated, in the section on developing a method of negative justification, that a violation of the preferences of some entity via the citation of preferences one shares with one's patient is, *ceteris paribus*, unjustifiable. The *ceteris paribus* condition, here, is indicative of the assumption that the PGC does not explicitly prescribe that we act against a being's preferences in order to fulfill some duty prescribed by the PGC. If the PGC *does* prescribe that we act against some commonly held preference, it constitutes a sufficient reason for doing so, since the citation of preferences, alone, is an invalid basis for justification. However, as we have demonstrated before, without the PGC's prescription to the contrary, it is still unjustifiable to act with the implicit assertion, acting as a sufficient reason, that one acts against one's own preferences. This, however, is shaky ground, since we have already demonstrated that preferences are unjustifiable bases of judgment. This citation of Hare's theory in the context of animal rights, however, is for the sole purpose of demonstrating that one may at least not justify the harming of an animal via the citation of one's own preferences to do so, where one's preferences constitute commonly held preferences between the two parties. This account provides what might be considered a negative justificatory basis for animal ethics; in other words, it might provide us with a fairly concrete and reliable account of ways in which we may *not* act toward non-agents. While this method is potentially very interesting, I shall not explore it here, for my main purpose is to elucidate the consequences of Gewirth's dialectical approach. As such, I defer to other theorists to further develop any potential methods based on the commonly-held-preferences interpretation of Harean universal prescriptivism.

Let us now trace the conclusions of Gewirthian theory and see where they lead with respect to animal ethics. Although an entity has rights by virtue of its being an agent, animals in general have rights, Gewirth claims, according to the degree to which they approach having agency. Many animals, for instance, exhibit purposive behavior and can, thus, be said to have some form of intentionality. Thus, animals,

he argues, possess the generic rights in a degree proportional to the degree to which they approach having agency; this is Gewirth's version of what is called the Principle of Proportionality. Many animals, including whales, dolphins, octopi, and bonobos, exhibit certain cognitive capacities that are relevantly similar to ours, with whales and dolphins, specifically, often even exhibiting intentioned resistance to captivity.⁷³ By "relevantly similar," I mean that these similarities are morally relevant such that if I decided, say, to kill a dolphin, I would be implicitly asserting, "It is not the case that anything that exists which has intentionality ought not to be killed." This, however, is highly problematic, since agents also have intentionality. If we have a duty not to kill beings that have intentionality, it follows that those beings have rights against us not to be killed. Certainly, there will be dilemmas in which one has a choice between violating the rights of an animal and violating *the same rights* belonging to an agent, but agent rights have priority over those of other animals in such situations. If there is a choice between violating *different rights* of an agent and a non-agent, say, for example, the animal's right not to be killed and the agent's right not to be harangued, we might well choose to violate the agent's right. "But wait a second," one might object, "I thought you said that you would be contradicting yourself if you violated the right of an agent." The point of dilemmas, however, is that one is generally *forced* to make a decision between two violations of rights or, otherwise, two unjustifiable choices. In this case, Gewirth's theory seems to justify haranguing the agent, but this is not entirely clear. As we shall see, there seem to be certain limitations in Gewirthian theory's applications to animal ethics, limitations that I believe can be remedied only by a more fundamental virtue-theoretical grounding.

While the aforementioned account is what Gewirth himself suggests derives from his theory, Beyleveld contests Gewirth's interpretation.⁷⁴ Beyleveld, I believe quite rightly, argues that such an interpretation of the Principle of Proportionality does not logically follow from Gewirth's theory because, on Gewirth's theory, an entity possesses agent rights *only* by virtue of being an agent in the full sense. Let us see the consequences of this. If, per the above example, I kill a dolphin, I do indeed assert "It is not the case that anything that exists which has intentionality ought not to be killed." On Gewirth's theory, however, this is irrelevant; I am not at all, as Gewirth himself seems to imply, contradicting myself. This is because, on Gewirth's theory, such a statement as "I must not kill X," or in other words "I have a duty not to kill X," derive from a correlative right claim made on the part of other agents which I cannot deny without denying my own right claim and thus my agency itself. Thus, there are no rights which derive from mere intentionality, that

⁷³ Traci Warkentin, "Whale Agency: Affordances and Acts of Resistance in Captive Environments," *Animals and Agency: An Interdisciplinary Exploration*. Ed. Sarah E. McFarland and Ryan Hediger. (Boston: Brill. 2009): 23–45.

⁷⁴ Beyleveld and Roger Brownsword, *Human Dignity in Bioethics and Biolaw*. (Oxford: Oxford University Press. 2001); Beyleveld and Shaun Pattinson, "Precautionary Reasoning as a Link to Moral Action." *Medical Ethics*. Ed. Michael Boylan. (Upper Saddle River: Prentice-Hall. 2000): 39–53.

is, intentionality without the accompaniment of the assertion of right claims. If we re-examine Bandura's social cognitive theory, we might understand this point a bit better. Agents assert rights due to their agency. Such an assertion of rights, however, is due to a specific sort of agency, namely the multi-tiered model laid out in Bandura's theory. These right claims, in other words, derive from a form of agency which includes a capacity for self-reflection, the establishment of self-standards and self-efficacy judgments, meta-level motivation, and a complex self-system. This is the only way in which right claims can be asserted at all, either implicitly or explicitly. Thus, on this account, non-agents have no rights whatsoever. This, to be clear, includes all so-called "marginal agents," such as infants and those with certain severe mental disabilities. In order to steer away from what would rightly be considered potentially genocidal conclusions, Beyleveld and his colleagues have developed what they call the Precautionary Principle, which states:

If there is no way of knowing whether or not X has property P, then, in so far as it is possible to do so, X must be assumed *to have* property P if the consequences of erring in presuming that X *does not have* P are worse than those of erring in presuming that X *has* P (and X is assumed *not to have* P if the consequences of erring in presuming that X *has* P are worse than those of assuming that X *does not have* P).⁷⁵

From this, it accordingly follows that:

Although X is apparently only a partial agent, *because the proposition that another creature is an agent is a metaphysical one and human reason is limited in such matters*, I cannot infer that X *is not* an agent. Just as I cannot *know with certainty* that X *is* an agent when X is an ostensible agent, so I cannot *know with certainty* that X *is not* an agent when is apparently only a partial agent. So, even though X is apparently only a partial agent, there remains a risk that if I suppose that X is not an agent, and act accordingly, X is an agent, and I will have deprived X of the protection of the PGC to which it is categorically entitled.

This, then, justifies his reformulation of Gewirth's original interpretation of the Principle of Proportionality. Instead of the proportionality's holding between the non-agent's rights and the degree to which a non-agent approaches being an agent, Beyleveld's argument leads to the conclusion that "the moral status of creatures is *proportional* to the probability that they are agents."⁷⁶ This is admittedly a more satisfying account than that which Gewirth himself has offered. I believe, however, that Beyleveld's account remains problematic. It might seem to Beyleveld and his colleagues a matter of splitting hairs, but the following point is, I believe, important: although, in practice, we do end up according rights to marginal agents, this is, as it were, a fluke. In fact, according to Gewirth's theory, if we had full and perfect knowledge (i.e., absolute certainty) that certain entities were indeed non-agents, the theory would not accord as immoral or unethical in any way the mass killing of such marginal agents. We must not, however, stop there; mass torture, mass enslavement, and mass experimentation would be perfectly legitimate as well. Indeed, there seems even to be a potential problem in Beyleveld's account with respect to

⁷⁵ Beyleveld and Brownsword, *Human Dignity*, 122.

⁷⁶ *Ibid.*, 123.

the opportunity costs of action with respect to apparent non-agents. He states: “It is. . . under the PGC, worse to err in assuming that X is not an agent than to err in assuming that X is an agent. Thus, *for practical purposes*, precautionary reasoning requires that X be assumed to be an agent.”⁷⁷ This, however, seems problematic, for there is an opportunity cost inherent in every decision. While Beyleveld claims that it is, under the PGC, worse to err in assuming that X is not an agent, this does not, upon closer, inspection, appear to be the case. On the contrary, would it not be worse, in many cases, to err in assuming that X *is* an agent? Let us take the following example: John decides that it is a very worthwhile use of his time to volunteer with an organization that rescues animals from abuse and neglect, perhaps the ASPCA. John, however, has just learned of Gewirth’s argument to the PGC, has accepted it, and has decided to implement it in guiding his decisions. He realizes, however, that Gewirth’s theory implies that, given the choice between assuming that X (which is most likely not an agent) is an agent, and thereby deciding to attempt to provide for its needs and so forth, and assuming that Y (which *is* most likely an agent) is an agent, and thereby deciding to attempt to provide for its needs and so forth, John should choose to give aid to Y. In other words, as John realizes, all actions incur an opportunity cost, and so John must not run the risk of assuming that X is an agent when it most likely is not. It follows from this that, instead of volunteering with the ASPCA, John should volunteer at another organization aimed, perhaps, at aiding people in highly impoverished countries. John makes the further realization, again on the PGC, that his time would be best spent convincing all others to do the same, in other words to stop giving aid to probable non-agents and to begin devoting this time to aiding probable agents. Only once we have maximally upheld the rights of agents (and thus solved all salient global issues), he concludes, may we ever justify aiding a non-agent. In essence, every decision, we must see, is a normative dilemma, merely by virtue of its having an opportunity cost, and so, because of this, we must follow Gewirth’s theory where it logically leads, namely to the total neglect of probable non-agents in favor of probable agents. I of course could appeal to what I perceive to be people’s general intuitions concerning this kind of conclusion as being incorrect in some way, but I will not, as I have before promised, adopt what I consider to be this flimsy mode of argumentation. Instead, what I am attempting to make salient here is not, strictly speaking, that Gewirth’s theory has seemingly bad conclusions with regards to marginal agents but rather that, in a very vital sense, Gewirth’s theory does not make any firm conclusions about how we are to treat marginal agents at all, besides neglecting them or favoring probable agents as opposed to them. Thus, my essential claim here is that Gewirth’s theory contains gaps in its applications as to how we are actually to treat non-agents when we do inevitably come to interact with them.

I believe, however, that virtue theory provides a solution to this problem. By grounding the moral status of non-agents in a virtue-theoretical framework, we see, I believe, that Gewirth’s theory can ultimately survive this critique. One way we

⁷⁷ Ibid., 128.

could justify the moral status of marginal agents is through the use of a Footian, pseudo-Thomist, interpretation of virtue ethics. According to this interpretation, all living things, including plants and trees, have natural goodness or value simply by virtue of their being alive. Strictly speaking, virtue ethics, and this interpretation is included, does not establish any rights or duties; it establishes virtues. On this interpretation, we would not speak of agent rights or duties or animal rights. Instead, we would speak simply of agent and animal virtues and goods. Thus, it would be virtuous to act in a certain way toward living things in general, and it would be vicious to act otherwise. In other words, one would degrade and vitiate oneself as an agent if one were to act in a certain way toward animals. This approach, we see, is extremely far-reaching, for it establishes a mode of treating nature in general, and not just a mode of treating animal life. On this interpretation, even non-sentient life is taken into account, and this is truly fascinating. While fascinating and indeed very insightful, this view suffers from salient problems as well. We will explore these problems, as well as Foot's general account of virtue ethics, in Chap. 5. For now, without establishing the actual content that grounds a justification of animal moral status (which will come in later chapters), we will explore the merely formal elements through which I claim the problem of Gewirth's theory regarding marginal agents can be resolved. It is only when we arrive at the final stages of my theory's elucidation that we will see clearly why and how such an approach works. In the meantime, therefore, the reader will have to excuse the use of what might seem to be vague and ill-defined terms such as "*telos*," "debasement," "worse functioning," "evil," and "best rationality," which will be defined and characterized *in extensio* in later chapters.

Once again, as I have been repeatedly emphasizing, we will need to more deeply analyze the concept of experiential rationality. It is from the unified concept of logical and experiential rationality, taken together, I will later demonstrate, that the normative legitimacy of the concept of individual *telos* derives. Due to the inherent value of the elements that underlie the structure of *telos*, namely the essential elements of rationality, it would be evil to engage in any action which led to the debasement or otherwise worse functioning of such elements. Such elements, however, are not the only things which possess value; underlying the value of these elements of rationality is the necessary value of things which make rationality possible or which, in other words, make the instantiation of rationality in reality more probable; such things include life, as well as existence more generally. Now, these things by themselves are not in any way certain to yield rationality, but, by making the instantiation of rationality in reality more probable, they gain their value. Animals and marginal agents more generally, then, attain their value in a way somewhat similar to how Gewirth originally conceived the issue, namely as a proportionality between the rights of the entity and the degree to which they approach having agency. My reconstruction of the Principle of Proportionality, however, is substantively different. Instead of framing the principle as a proportionality between "rights" and the "degree" to which an entity approaches having "agency," I choose to state the principle in the following way, as derives directly from the premises of my virtue-theoretical account:

The Principle of Proportionality: *The value of a non-rational entity is proportional to its capacity for rationality, as measured by the probability that it will attain rationality, given the conditions and limitations of its physical embodiment.*

What this implies, for instance, is that that animal has greater value for whom the probability is greater that, given an alteration of its brain physiology, genetic make-up, or even cybernetic make-up, it would become rational. It does not matter whether the difference is between a probability of .000000045 and a probability of .000000046; it is merely a matter of what, in principle and regardless of energy and so forth committed to the task, is the more probable event. Thus, this is a radically different result than that offered by either Gewirth or Beyleveld, although it bears a certain family resemblance to their respective conclusions. While again this argument might seem in some degree ambiguous, due as I stated to its being a merely formal argument without the determinacy provided by theoretical content, my argument for a post-dialectical aretaic model will later make abundantly clear that such an account demonstrates that animals and other marginal agents do indeed have inherent value and, importantly, rights that derive from such value. There exist, therefore, concrete animal rights.

3.3.4 Applications to Positive Rights

We shall now address the second question: “What about positive agent rights?” As Gewirth and Beyleveld have demonstrated, all agents must *at least* not interfere with the rights to freedom and well-being of agents. But is there anything beyond non-interference? What about *helping* other agents to exercise their rights to freedom and well-being? After setting out the negative account of rights and duties prescribed by the PGC, Gewirth revisits an earlier premise, namely “I must have freedom and well-being” or, to state it more clearly, “I must have the freedom and well-being to pursue purpose E.” While this entails “Others must not interfere with my pursuit of purpose E,” it also entails, he claims, “If I cannot pursue purpose E on my own efforts due to my not having the freedom and well-being adequate to pursue it, others must help me to acquire the freedom and well-being adequate for me to pursue it.” In other words, “I must have X” entails both a negative and a positive statement, namely “Others must not interfere with my having X” and “Others must help me to have X if I cannot have it on my own efforts.” “X,” as I have slightly reformulated it, is intended to stand in for “freedom.” If “I must have the freedom and well-being to pursue purpose E” is entailed by all voluntary action, and if both negative and positive claim-rights are entailed by this judgment, then it follows from Gewirth’s theory that there exist both negative and positive agent rights. It might be even more edifying to think about the deduction of positive rights from this statement in the following manner: “I must have X” entails two kinds of judgments, namely what I will refer to as an indicative judgment and a conditional

judgment. The indicative judgment takes the form "Others must not interfere with my having the freedom to X," while the conditional judgment takes the form "If I did not have the freedom to X, then others would have to aid me in obtaining the freedom to X." From this reconstruction, we can see that common to both judgments is the acknowledgment of some present state of affairs, namely the possession of the freedom to X, which is posited as ethically required. From the positing of such an ethical requirement concerning the present follows some judgment about what specific conditions would therefore have to be met if the freedom to X were absent in the present, namely those conditions which lead to the fulfillment of the freedom to X. This is the logic upon which rests Gewirth's derivation of positive rights, and I believe it to be a sound one. Of course, Gewirth, Beyleveld, and others have already undertaken extensive analyses to demonstrate the soundness of the position. One fatal assumption which Gewirth makes, however, which is common to most proponents of positive rights, is that the formation of a so-called supportive state derives from the existence of positive rights. It does not, however, follow from the PGC, as Gewirth suggests, that the supportive state, or even a state at all, ought to be established; such a claim is one concerning an empirically contingent feature of society, the State, that bears relevance only to the modern era (i.e., the only era in which such an institution as the State has thus far existed).⁷⁸ I believe, rather, that all Gewirth can possibly derive from his ethical theory, as it stands, is that some supportive institutional social structures are ethically required. In other words, nowhere in Gewirth's theory is it entailed that a central power with monopoly-force rights, fixed borders, and so on, ought to be established or ought to exist; all it entails is the minimalist thesis that individuals ought to establish supportive structures. If we follow this minimalist thesis to its logical conclusion, namely by pointing out that no such thing as a State is logically entailed by the PGC, I believe we in fact arrive at a *de facto* libertarian position, one that demands a burden of proof for the legitimacy of any and all limitations on freedom. Thus, there are many unwarranted assumptions inherent in Gewirth's political theory, but we will not be able to do justice to the theory by examining it here, along with the elements of his ethical theory. We will, however, somewhat briefly revisit the topic of political theory as we progress in the overall analysis. Although a comprehensive analysis of political theory is beyond the scope of this book, I will elucidate some of the immediate implications of the dialectical model for political theory when I explore Habermas's dialectical approach.

⁷⁸The only historical exception to this, if there is one, would be the ancient Roman Empire.

3.3.5 *Applications to Classic Moral Dilemmas, the Principle of Double Effect, and Intention*

We will now address the second question concerning moral dilemmas and self-defense. Let us take up self-defense first. Self-defense is not usually considered a type of moral dilemma, for all too often it seems to be assumed among non-pacifist philosophers that self-defense is a straightforwardly justifiable act. The concept of self-defense is, however, more complex than this. In recent years, Eric Reitan and others have acknowledged the dilemmas inherent in self-defense, and Reitan has specifically addressed Gewirth on this subject.⁷⁹ According to Reitan, Gewirth's position on self-defense, namely that killing in self-defense is justifiable if one's life is in imminent danger, is inconsistent with and does not derive from the PGC. Reitan's analysis, I will demonstrate, is flawed because it fails to take into consideration a key deontological implication of the PGC: that what matters in dilemmas is intention and the assertions or reasons one must propose as one's intention. I am not proposing that consequences do not at all matter; this is a point with which even Kant would agree. Rather, I am contrasting my own Gewirthian position against simple consequentialism. In a situation in which a murderer imminently threatens my life, simple consequentialists would have to admit that, whichever course of action I choose, I am justified in my decision because the consequences are the same: one person dies or one person dies. What I am arguing, along with nearly all deontologists, is that there is something more to such situations and that this something is intention. When I act, I must have certain intentions, and, make no mistake, these are intentions *concerning the consequences*.

Suppose that someone is advancing toward me with a gun with the intention of shooting and killing me. To the best of my ability, at least, this is what I am able to judge. My assailant, I judge, will kill me if I do not do something, namely kill him. But does not the PGC proscribe my killing other people? It seems that, whichever course I take, I will either maximally violate my assailant's right to freedom or have my own rights maximally violated. Whichever course I choose, I have violated the PGC, right? Not quite. Let us examine this more closely. First, let us examine what, on Gewirth's theory, would be the morally relevant criteria entailed by the situation.

Morally Relevant Criteria:

1. *The PGC states that both I and my assailant have equal rights to freedom.*
2. *My assailant is attempting to do something proscribed by the PGC, namely maximally violate my right to freedom.*

⁷⁹ Judith Jarvis Thomson, *The Realm of Rights*, (Cambridge: Harvard University Press. 1990); Eric Reitan, "Self-Defense and the Principle of Generic Consistency: Must Gewirth Be an Absolute Pacifist?" *Social Theory and Practice* 32, no. 3 (2006): 415–438; Suzanne Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide*, (Cambridge: Cambridge University Press. 1994).

3. *In order to prevent my murder and thus uphold my right to freedom, it would be necessary to maximally violate my assailant's right to freedom.*

Therefore, all possible choices I can make in this situation entail the maximal violation of an agent's right to freedom. However, I *must* make one of the two choices; I cannot escape making a choice. According to the PGC, I have a duty to do only what is necessary to uphold my freedom without violating another agent's right to freedom, but this is not an option in this situation. So, I must have some discriminating criterion to guide me in choosing which agent's maximal right to freedom it is more justifiable for me to violate. The discriminating criterion can be found, in part, in the previous statement: "I have a right to freedom, and I have a duty to do *only* what is necessary to uphold my right to freedom." I emphasize *only* what is necessary because the PGC does not prescribe any more than what is necessary; indeed, it proscribes more than what is necessary, that is, what is a necessary condition of upholding one's rights.

For contrast, let us examine my assailant's perspective. My assailant also has a right to freedom and has a duty to do only what is necessary to protect his freedom, but my assailant has a certain set of choices before him, namely:

1. *To kill me and thus violate the PGC by maximally violating an agent's right to freedom; and*
2. *To not-kill me and thus abide by the PGC by not interfering with an agent's right to freedom*

Therefore, only one possible choice on my assailant's part entails the maximal violation of an agent's right to freedom, and the other entails the refraining of interference with an agent's right to freedom. His choices, therefore, are not limited only to seemingly unjustifiable ones; he is in a position to make a justifiable choice. By attempting to kill me, therefore, he is not doing *only* what is necessary to uphold his right to freedom. What, we might ask, would be necessary for our assailant to do in order to uphold his rights to freedom without violating others' rights to freedom? It would be necessary, of course, to refrain from killing, among other things. Therefore, we can see that our assailant is clearly morally culpable for the constraining of my choices to two maximal violations of agent rights. Let us examine the assailant's choices more closely. By choosing to murder me, my assailant asserts:

1. *"It is better to maximally violate an agent's right to freedom."*
 - 1a. *"I ought to maximally violate an agent's right to freedom."*

From the PGC, we conclude that 1 and 1a are unjustifiable.

By choosing not to murder me, my assailant asserts:

2. *"It is better not to maximally violate an agent's right to freedom."*
 - 2a. *"I ought not to maximally violate an agent's right to freedom."*

From the PGC, we conclude that 2 and 2a are justifiable, obligatory.

We see, therefore, exactly how our assailant's choices are not constrained. To see this more clearly in contrast to my choices, let us examine my own choices more closely. By choosing to let my assailant kill me, I assert the following:

1. *"It is better for my assailant to maximally violate an agent's right to freedom."*
 - 1a. *"My assailant ought to maximally violate an agent's right to freedom."*
2. *"It is better for me to allow a maximal violation of an agent's right."* (my own)
 - 2a. *"I ought to allow an agent's right to be maximally violated."* (my own)

From the PGC, we conclude that both 1 and 1a and 2 and 2a are unjustifiable.

By choosing to maximally violate my assailant's right to freedom, I assert the following:

4. *"It is better to violate my assailant's right to freedom."*
 - 4a. *"I ought to violate my assailant's right to freedom."*
5. *"It is better to uphold my freedom; I acknowledge my duty to uphold my freedom."*
 - 5a. *"I must uphold my right to freedom by doing only what is necessary to do so."*

The reader should begin to see the logical connections here. While my assailant can choose not to maximally violate my right to freedom, he chooses the opposite and does something unjustifiable. By contrast, *both* of my choices entail a maximal violation of an agent's right to freedom, *but* I must still determine the option, if any exists, which entails the upholding of *some* right to freedom. Can we find such an option? If we examine both of my possible choices above, we see that my first choice, to allow the assailant to violate my right to freedom, entails two unjustifiable assertions (1a and 2a). By choosing to violate my assailant's right to freedom, there are, yet again, two entailed statements. One of the statements entails a violation of an agent's right to freedom (4a), but the other statement entails an upholding of an agent's right to freedom, namely my own. Thus we see yet another application for which deontology proves superior. The consequences, in terms of utility, are the same whichever choice I make: If I choose not to defend myself, I die and the assailant lives, and if I choose to defend myself, I live and my assailant dies. Quantitatively, in terms of utils or some such other fiction, the consequences are the same. It is, however, the *intention* here that is morally relevant; that which is morally relevant, in other words, is that which I can consistently advocate. I cannot, in any sense, consistently advocate 1a nor 2a, but I *can* consistently advocate my killing of my assailant, which entails at least one justifiable assertion, that is, at least one assertion that is in accord with the PGC. One might object that we might also choose to allow ourselves to be killed and, thus, advocate an upholding of our assailant's right to freedom. This objection, then, highlights the necessity of emphasizing an important point about intention: It is not only my discovery of some choice that is superficially in accord with the PGC that is important, that is, one which entails the upholding of some agent's right to freedom; rather, my discovery of some choice to uphold some agent's right to freedom, when faced with such a dilemma as presented above, must *not* be accompanied by an implicit approval (i.e., an "ought"-judgment) of the agent's

unjustifiable choice. *If my adoption of a choice entails such implicit approval, then I am violating the PGC because I am accepting as true a false moral assertion.* Whereas my choice to spare the assailant's life yet terminate my own would entail the judgment, "Agents have the right to the freedom to kill each other," my choice to spare my life yet terminate the assailant's entails the judgment, "Agents have the duty to do *only* what is necessary to uphold their right to freedom." Of the two choices, it is clear *from my intention alone and not from the consequences*, that my choice to kill my assailant and thus spare my life is the ethically justifiable, although regrettable, choice.

One might press this example further, as Thomas Scanlon does, by pointing to the case of an attack on the part of a person with some sort of temporary incapacity for acting voluntarily. In such a case, however, intention still matters. Of course, we may take it as a given that the attacker is not making any voluntary judgments and thus cannot be said to have intentions for acting as she does, but this does not nullify intention. If the person being attacked can make judgments and act voluntarily, then it is hardly relevant that the attacker cannot do so. As with the previous example, we can examine the intentions of the person being attacked and determine which assertion or set of assertions is justifiable. Whichever assertion is justifiable, or consistent with the PGC, is that which the person being attacked must thus follow, if he can.

We may now proceed to the second part of the last question, that concerning moral dilemmas. Gewirth and Beyleveld have addressed, in some depth, moral dilemmas and the PGC's applicability to them, but, as I will demonstrate, the PGC yields far-reaching solutions to common, seemingly intractable dilemmas that neither theorist has directly addressed. Let us explore, in order, some classic moral dilemmas as Hauser has presented them in *Moral Minds*, as his presentation is highly instructive. Hauser's presentation demonstrates how difficult, on various theories, it is to keep consistent with one's judgments.

First is the classic trolley dilemma, with which all philosophers are familiar. You are standing on an overpass of some sort that overlooks a trolley track that forks at a certain point. On one fork of the track stand five people, and on another fork of the track stands one person. The trolley is coming, and it is headed for the five people, but you have access to a switch that can make the train transfer to the other track instead. So what do you decide? Most people, Hauser explains, have the instinct to flip the switch, thus sparing the five people and resulting in the death of the one person. What principle have you chosen to make this decision? Perhaps that of utility?

The second is the modified trolley dilemma, which is also sometimes called the bridge case. You are again standing on an overpass, but this time there is no fork. Instead, it is the five people alone that are on a track. This time, you do not have access to a switch; instead, there is a very large man on the overpass next to you who, if dropped onto the tracks in front of the moving trolley, would stop it and, thus, save the five people.⁸⁰ Once again, the trolley is coming, and you must make a

⁸⁰ We must set aside the intrinsic implausibility of this situation as it applies to physical laws of momentum and energy conservation. Let us, for instance, believe that this man is extremely massive.

choice: push the man over the overpass and kill him to save the five people, or let the five people die. If you are using your principle of utility, you must, despite what Hare might plead with regards to the improbability of the situation, push the man over and kill him to save the five people. Have you still kept your principles consistent?

The third dilemma is the well-known organ-and-tissue-donation dilemma. This time, there are five people in a hospital who are in dire need of organs and tissues without which, in a short time, they will die. Some person suddenly finds her way into the hospital for the treatment of some condition she has. Somehow, the physicians who are in charge of caring for the five patients in need of organs and tissues discover that this person is organ-and-tissue-compatible with all five patients; it is a miracle! Or is it? Remember, these are vital organs and tissues; if she loses these, she will die. So, you face a choice: kill this person and save the five people, or refrain from killing her and let the five die. Have you still kept your principles consistent?

Let us see what the PGC has to say. In all three scenarios, there is a conflict of rights and a conflict of duties: If I refrain from violating the rights of one group, I end up violating the rights of another group; if I fulfill my duty to one group, I fail to fulfill my duty to another group. Add to this that these are the same rights; there is no possibility of judging, using “degrees of necessity” as Gewirth calls it, which right or duty takes priority. Remember also that, although the PGC dictates that we help, and in this situation rescue, all agents we are in a position to rescue without comparable harm to ourselves, we are not in a position to do so. In other words, the PGC does not prescribe a rule of utility; it only prescribes that one save all agents, *not as many agents*, that one is in a position to save. I am, however, in the position to save only one party or the other. What, then, is the solution? In Stephen Nathanson’s article, “Can Terrorism Be Morally Justified?,” he argues against the principle of double effect on the grounds that it is fatally weak. Its form, “I may effect evil A in order to effect good B only if I do not intend to effect evil A,” contains the expression “I do not intend.” This, he states, is the problem. If, in order to stop the invasion of some evil military force, it is necessary to bomb some specific site, he states, the principle of double effect offers the bombers an unjustified moral excuse for killing innocents. According to Nathanson, we need a new principle, what he calls the Bend-Over-Backwards Principle.⁸¹ Instead of “I do not intend,” this principle takes the form of “I intend not.” In order to justify such a bombing, Nathanson states, even if there is an imminent threat of invasion, the bombers must “bend over backwards” and do absolutely everything they possibly can not to kill a single person in the process of their bombing the site. The Bend-Over-Backwards Principle, or what I shall call the Contrary Double-Effect Principle,⁸² is, as it turns out, actually prescribed by the PGC in its implicit clause “Do only that which fulfills the necessary condition of A without

⁸¹ Stephen Nathanson, “Can Terrorism Be Morally Justified?” *Social Ethics: Morality and Social Policy*. Ed. Thomas A. Mappes and Jane S. Zembaty. New York: McGraw Hill, 2007. 322–332.

⁸² I find the “Contrary Double-Effect Principle” a more accurate label, for it emphasizes that this principle and the Double-Effect Principle are, in terms of formal logic, contraries of each other.

violating the rights of *any* agents." We must, therefore, do everything in our power *not* to do anything that results in the death of anyone; we must, therefore, *intend not* to have anyone die.

Let us first re-examine the latter two scenarios. One main factor that sets these two scenarios apart from the first scenario is that, in the first scenario, both sets of patients cannot exercise choice; they do not have the ability to make a choice about the situation. In some very real sense, they cannot act as agents. In the latter two scenarios, by contrast, it is not the case that both of my patients cannot act as agents; one group can exercise choice in the situation, and another group cannot. This difference between the two situations is vital, and for this reason: If the patient is capable of exercising choice, and especially if his choice is a justifiable one, we are not justified in forcing him to do something unjustifiable. This sort of peculiarity does not exist in the first scenario. Let us, then, examine the two patients' set of choices in the latter two scenarios. In the second scenario, this very massive person has either to choose (1) to jump in front of the trolley and kill himself to save the five; or (2) to let the five die and spare himself. In both instances, he is making a conscious decision for which he is fully aware of the consequences. One might conclude, then, that one cannot, thus, say that he *intends not* either outcome. One might argue, for instance, that through the exercise of either choice, since he is fully aware of his choice's consequences, he intends those consequences. Although we might say, for instance, that he wishes very intensely that the people before him would not die if he spared his life, we will have to admit that he intends the consequences, for he must still state or assent to the following: "It is better for X to die than for me to die" and "I ought to live." The point being made here is that "I ought to live" is actually a necessary condition of the others' dying. As Gewirth points out in another rescue scenario, the non-agentive factor, be it a trolley, a falling piano, or water which drowns someone, is only a necessary condition of the bad upshot if an agent who is able to counteract the non-agentive factor is *not* present. Once such an agent is present, the agent's choice not to counteract the non-agentive factor *becomes* a necessary condition, and the non-agentive factor becomes a sufficient condition. Thus, the very large person in the second scenario *must* intend either outcome, yet he must still make a choice. We shall see if this sort of interpretation of intention holds. This conception of intention, I shall conclude, is incoherent. As we examine the following positions, we shall see why.

There are three seemingly plausible positions that we could, at this point, derive from the PGC. We shall begin with what I will call Position A.

3.3.6 Position A: Incompleteness and Equal Justifiability

For the man standing on the bridge, both choices are equally unjustifiable, and this also means that both choices are equally *justifiable*. This indeed applies to both latter scenarios. At this point, it is important to emphasize, if it is not already obvious, that these scenarios are *ex hypothesi*; there are no other choices available than those which

I have given. Hare would tell us that these scenarios, then, are not ones which could actually happen in real life and that, therefore, we are wasting our time in discussing them. On the contrary, even scenarios that are highly unlikely are important to discuss, since, if our principles and our overall model are sound, then they should be capable of being shown to be applicable in even the most unlikely of circumstances. Let us, then, examine what the very massive person should do. This person might decide that utility is relevant here, but the PGC does not prescribe any such principle. Since both choices are equally justifiable, the final criterion that the man must use is one that is wholly a personal one; since both choices entail a contradiction, literally *any* justifying reason follows. Certainly, under more likely circumstances, this person would probably have a greater amount of criteria from which to make his choice. He might, for instance, have a family or a partner to which he has duties. He might also realize that he could simply yell to these people to get them off the track. He might have any number of alternative choices, but if the only choices available to him are the two stated under the *ex hypothesi* conditions, either choice on his part is justifiable. This has important consequences. In the latter two scenarios, I am also obviously an agent; I must make a choice. In both scenarios, yet again, I have two choices: kill one to save five, or fail to save five and let one live. Once again, the Contrary Double-Effect Principle, entailed by the PGC, mandates that I do everything I possibly can not to perform an action that results in a maximal violation of an agent's right to freedom. Both of my choices, however, entail a maximal violation of an agent's right to freedom. Here is where my patient's ability to make a justifiable choice becomes relevant. In order to see this, let us examine what assertions I implicitly make in both of my choices.

1. *"I ought to do something which results in the maximal violation of five patients' (also agents) rights to freedom."*
 - 1a. *"I ought to maximally violate an agent's right to freedom."*
2. *"I ought to do something which results in the maximal violation of one patient's right to freedom, and my patient has the right to make whichever choice he wants, since both are equally justifiable."*
 - 2a. *"I ought to maximally violate an agent's right to freedom, and I ought to violate an agent's right to make a justifiable choice."*

Thus, although my choice to save the five patients by killing the one patient would result in the upholding of their rights to freedom, I must recognize that my other patient in this scenario, the man on the bridge, is also an agent. In other words, since he too must evaluate the scenario and make a choice about it, I must respond to his choices as I would any other agent. Since either choice he can make in this scenario is equally justifiable, it would be unjustifiable for me to violate his freedom to make such a choice. As the PGC implies, I must do only what is necessary to uphold agents' rights to freedom without violating agents' rights to freedom. Since my patient has the right to make either choice, and since they both entail the maximal violation of agent rights to freedom, I must recognize this choice as a right, a right that the other five patients do not have. If they somehow did have this right, then I would be even less justified in killing either

party. The most I could justifiably do is facilitate a discussion or agreement between the two parties.

In the first trolley scenario, I take on a relevantly similar role as my patients in the latter two scenarios. In other words, I must make one of two equally unjustifiable decisions: let one agent die or let five agents die. Most accounts of this scenario use the Double-Effect Principle to attempt to justify the priority of the five over the one because, problematically, I *do not intend* the one to die. Since Nathanson, supported additionally by the PGC, has demonstrated this principle to be inadequate, I must admit that I must *intend not* to have either party die; I must do everything in my power not to have either party die. According to the objection surveyed previously, however, I cannot do this; I must intend that one party rather than another die. Again, since, *ex hypothesi*, all I have are two unjustifiable decisions, I may make either decision with impunity, although Hauser would point out that most of us would choose to flip the switch. Under more realistic circumstances, of course, we could flip the switch and yell to the one person to get off the track; flipping the switch is a sheer act of practicality, for one person will probably be able to get off the track more quickly than five.

3.3.7 Position B: The Self-Contract

We, once again, realize that we face two unjustifiable decisions in both case 1 and 2. There are, however, some relevant differences between the two cases. Once again, we shall begin by taking a look at what makes case 2 particularly peculiar. In this case, we realize that one patient, namely the large man above the tracks, is an agent and can act as an agent; he needs, as it were, no surrogate. Our other patients are agents but cannot act as agents since they have absolutely no choice over what happens in the situation and no ability to aid me in facilitating my or the large man's choice. Therefore, we have two choices, presented as two claims:

1. *I ought to maximally violate the rights to freedom of five agents.*
2. *I ought to maximally violate the rights to freedom of one agent, and I additionally ought to accomplish this by violating his right to exert his agency via choice.*

If we are following the PGC, the most logical conclusion to make is that it is worse for us to push the large man off the balcony to save the five people because it is equally bad to maximally violate agents' right to freedom, but it is worse to maximally violate an agent's right to freedom *and* violate his right to exert his agency via his own power of choice. One might object that it is justifiable to push the man over *because* his choice to save himself is unjustifiable, for instance, because (a) the other five or at least one of the five that will die are very young; (b) the other five or at least one of the other five are people whom you love and thus with whom you have a strong compact; (c) the principle of utility dictates so. While all of these seem like they might provide sound reasons for acting in this way, let us keep this case an *ex hypothesi* one in which none of these other factors are in play. In this *ex hypothesi* case, the prime reason why we ought not to push the man over is

due to what I shall call the Self-Contract. All persons have interests which they ascribe to some future self, and they have certain expectations concerning the fulfillment or at least partial fulfillment of these interests. The intentions toward the fulfillment of these interests are, as Bandura has shown, formed with regards to the ideal self formed by the agent's self-standards and self-efficacy judgments. These self-evaluations and subsequent perceived future interests are, as it were, a vital component of the cognitive phenomenon of free agency. They are, in effect, a vital expression of an individual's freedom; indeed, they are part of what Sen refers to as the opportunity, versus the process, aspect of freedom. This set of expectations is, in essence, a sort of contract with oneself of the following form: "I will and intend to attain, carry out, or effect X, and X will make me a better person." If we consider Parfit's theory of personal identity, we can perhaps understand how such a contract with oneself makes metaphysical sense. Because this large man has such a Self-Contract, you force him to maximally violate this contract by pushing him over. While it is true that those on the track have Self-Contracts as well and that I am effectively negating their Self-Contracts by allowing the large man to live, the vital difference, once again, is that the large man can act freely while the others cannot. In pushing him over, I must conclude not simply "It is better to violate the Self-Contract of an agent" but "It is better to violate the Self-Contract of an agent *who can choose* to break or uphold it; I must negate an agent's choice *he actually has* to uphold or forfeit his future interests." This distinction between choices can alternatively be seen, using Sen's terminology, as a difference between choosing to violate *either* the opportunity-aspect freedom of some individuals (the five people) *or* choosing to violate the opportunity-aspect freedom and the process-aspect freedom of one person (the large man).⁸³ In contrast, the five people need a surrogate; they cannot act freely, for they have no process-aspect freedom. Of course, it is extremely tragic that the other five die, but what the concept of the Self-Contract demonstrates is that it is a perfectly sound choice for the large man to make in this case to exert his existing agency to protect his future interests. Looking down upon the track, looking upon his entire life, and pondering all of his future interests, future goals, and future memories, it is a perfectly legitimate decision on his part to uphold his Self-Contract and all of this rich future agency. Since *ceteris paribus* this seems to be a perfectly legitimate decision, it cannot be justifiable to push the man over.

If we consider case 1, the situation is very different. In this case, both parties need a surrogate, for neither party has the power to make a choice about its fate. Thus, a criterion of choice is irrelevant here. If the criterion were relevant here, we would have to prioritize it above other principles such as that of utility. Such factors as the age of the parties involved or their personal relationships to you would also presumably have to be taken into consideration before utility, for they are derivable from concerns related to the deprivation of freedom. This, however, is a much longer discussion than cannot be substantively had here. What is clear, however, is

⁸³ Sen, *Rationality and Freedom; The Idea of Justice*.

that utility, if it has any role at all, is a criterion of last resort; the mere number of parties involved cannot trump these other factors related to the deprivation of freedom merely because the individuals comprising these parties are more numerous.

One might still object that cases 1 and 2 are not relevantly dissimilar. One might accuse me of gross inconsistency, in my presentation of position B. "How," one might ask, "could you possibly justify not pushing the large man over in the second case in order to preserve *more freedom* yet justify letting the one person die in the first case? Is it not the case that, by pushing the large man over in case 2, you actually substantively preserve *more freedom*? Is it not simply a fact that five people have more freedom than one?" While it is true that, in case 1, the five people individually have more freedom than the one person, so long as the case is considered *ex hypothesi*, all patients in the situation possess *exactly the same kind of freedom*, namely opportunity-aspect freedom. In the second case, however, we have a conflict: violate five people's opportunity-aspect freedom or violate one person's opportunity-aspect freedom *and* process-aspect freedom. Substantively, these are not the same types of freedoms and cannot simply be compared quantitatively via some utility function. Process-aspect freedom is the freedom to *choose*, and that is substantively more relevant to this kind of situation than merely *having* opportunity-aspect freedom with an accompanying Self-Contract. Why is this? This is because it cannot be overlooked as an empirical fact that the large man actually still has a choice and that, as a normative truth derived from the PGC, one cannot negate or violate another agent's choice in any situation unless that choice is unjustifiable and, importantly, unless one can be absolutely certain that such a choice is unjustifiable. As long as the patient has a choice (e.g.: he can act as an agent), then the nature of that choice must be considered before another agent can justifiably act. In effect, acting against the large man's choice in this situation entails a kind of inconsistency or absurdity: what distinguishes me as special or as having the authority to make a choice about the large man's fate as opposed to the large man himself? The answer is "nothing," for we are both agents with the capacity to make rational decisions, and, as Gewirth's Argument to the Sufficiency of Agency has demonstrated, it is the only criterion I may consider relevant in such a situation. Only a highly paternalistic ethics could allow us to push the large man over, although, let us be clear that, if it were shown that such a paternalistic ethics were sound, we would have to accept it. Due to Gewirth's theory, however, as well as the ideas of precaution and epistemological uncertainty inherent in both the dialogical approach and that of Beyleveld, both of which I shall analyze later, we see that such a paternalistic ethics is highly problematic. This explanation of anti-aggression is a bit more complex than Kant's maxim never to treat agents merely as a means to an end, but this reasoning is certainly in the spirit of this maxim.

One might press my argument yet further. "Is it in fact the case," one might ask, "that the large man's choice not to sacrifice himself for the five people is justifiable? Does he not have a duty to sacrifice his own life simply because a greater amount of freedom—the freedom of five people—is at stake?" The answer, put quite simply, is no. This man does not have any such duty to sacrifice his life for the five people

because the *types* of freedom that are at stake must be considered or prioritized before the simple *amount* of interpersonal freedom. What, you might ask, is the standard of qualitative comparison which determines this? Considering, once again, Sen's distinction between process-aspect and opportunity-aspect freedom, we might put this in different terms. "How many units of opportunity-aspect freedom," we might ask, "is equivalent to or greater than one unit of process-aspect freedom and one unit of opportunity-aspect freedom?" The simple answer, which entails a bit more complex reasoning, is that these two types of freedom cannot be measured against each other *in this way*, namely via a comparison of utilities. While Sen would have us believe that these two types of freedom cannot be measured against each other at all and, instead, constitute a sort of incompleteness, this is simply not the case. Although my response to the present objection will depart somewhat from Gewirthian theory, this will later prove instructive. Although we might not be able to compare process-aspect freedom and opportunity-aspect freedom *solely* against each other as some sorts of abstract entities (and, of course, we would never have to compare the two *by themselves* in real life), we can, indeed, compare them in a way that is more representative of what actually occurs in real life, namely as process-aspect freedom *and* opportunity-aspect freedom against opportunity-aspect freedom. For instance, we can conceive of someone who has the power of choice and the ability to plan and aspire (process-aspect and opportunity-aspect), someone who solely has the ability to plan and aspire, although she lacks the power of choice (opportunity-aspect), but we cannot conceive of someone who has the power of choice but who does not have the ability to plan or aspire, even if it consists in planning what to do in the next two seconds (process-aspect). As we shall see later on as we leave deontology to explore virtue theory, we shall see a bit more clearly that, while there is a large number of people on one track with opportunity-aspect freedom, there is no amount of this type of freedom that could count as quantitatively equivalent to one agent's opportunity-aspect freedom and process-aspect freedom; in other words, quantity, unless matched against something of the same type, can never be identical to quality. The possession of opportunity-aspect freedom and process-aspect freedom simultaneously just is an intrinsically more valuable, or better, state of affairs than the mere possession of opportunity-aspect freedom. Although the rule concerning quantity's lack of equivalence with quality does indeed derive from the PGC, it is this latter claim concerning the intrinsically greater value of the simultaneous possession of both types of freedom that will need to be further defended. We are not at the stage yet where it is appropriate to do so. More background and more explanation is first required. As we approach the subject of virtue theory, this position will be analyzed.

First, however, let us face another potential objection. "Can it not be said," one might ask, "that, by this standard, the man still ought to jump onto the tracks to save the five people? If the simultaneous possession of process-aspect and opportunity-aspect freedom is intrinsically better than opportunity-aspect freedom, then must we not maximize instances of both freedoms together (i.e.: save the five people so that they regain both of these freedoms)?" This might seem to be a damning objection, but let us consider the situation more closely. If the man jumped onto the tracks, he would eliminate his intrinsically better state of freedom, that is make himself worse than the five on the tracks, in order to grant each of them this intrinsically better state of

freedom. This is straightforward maximization of utilities, and this *would* be permissible, but, again, it would only be so if we were comparing two parties in relevantly similar circumstances. One cannot justify the effectuation of a quantitatively better outcome merely on the grounds that sacrificing one good yields five more of the same good. In order to justify such utilitarian-type acts, it is necessary, although not sufficient, that the goods at stake be of the same qualitative type.

3.3.8 *Position C: Double Effect, Contrary Double Effect, and Synthetic Double Effect Derived from the PGC*

Let us consider another possible alternative. Let us revisit the concept of intention. I stated earlier that Nathanson's Contrary Double-Effect Principle posed a problem for interpreting the moral dilemmas according to an intentionalist framework. I claimed earlier that, in case 2, the large man could not "intend not" either outcome because, on either choice, "he is making a conscious decision for which he is fully aware of the consequences." This notion of choice based merely upon full realization of the consequences is one interpretation of intention. The question, however, is whether this interpretation follows from the PGC, for we are justified in accepting it only if it so follows. If we interpret intention along similar lines as Anscombe has and incorporate Nathanson's reformulation of the Double-Effect Principle, however, we might find that simply making a choice while fully aware of that choice's consequences is not equivalent to intending those consequences. Instead, what we might find is that it is our aim or goal that is important, both in a negative and a positive sense. It is this understanding of intention, we shall see, as well as the Double-Effect Principle and its contrary, that actually derives from the PGC.

Let us examine each situation, but now specifically with regards to intention. Let us take case 1. Interpreting intention as "making a conscious effort to act in manner X by making X one's primary aim," we may analyze the statements entailed in case 1 as follows. From my perspective as the agent who must choose whether or not to flip the switch, I claim:

Double Effect: "*I do not intend*" (*I do not make a conscious effort to kill or let die—i.e., my primary aim is not to kill or let die/my primary aim is to let live*)

Contrary Double Effect: "*I intend not*" (*I make a conscious effort not to kill or let die—i.e., my primary aim is to not-kill or to not-let die/my primary aim is to let live*)

Since, in acting, I make my number-one aim not killing and not letting die, I both do not intend to kill or let die and intend not to kill or let die. This same interpretation can be applied to case 2. By pushing the large man over, I imply the following:

Double Effect: "*I intend*" (*I make a conscious effort to kill or let die—i.e., my primary aim is to kill in order to let live*)

Contrary Double Effect: *ditto*.

Clearly, on this somewhat Anscombean interpretation, as well as on the consequence-awareness interpretation, I do intend to kill. Most analyses of this scenario stop at this point, but, as we have been doing, let us once again imagine the statements of intention entailed by the large man's own choice to jump onto the tracks:

Double Effect: *"I intend" (I make a conscious effort to kill [myself] or let [myself] die—i.e., my primary aim is to kill [myself] in order to let live)*

Contrary Double Effect: *ditto*.

Even though on this interpretation the large man's choice and agency is not highly relevant to my own choice, we shall soon see how, in integrating the other positions A and B, this perspective is relevant. Before we conclude this position, let us first take a look at a scenario that, perhaps, might not survive the argument given in Position B. Let us consider a scenario, let us call it case 4, in which the five patients from case 3 are all somehow conscious but in need of vital organs without which they will shortly die but in which, this time, our prospective donor is a person in a coma with a high chance of recovery. One might criticize Position B on the grounds that it permits us to non-consensually remove the organs from the coma patient because "the five patients must not sacrifice their process-aspect freedom and opportunity-aspect freedom in order to save the process-aspect-deficient coma patient," but, once we analyze this situation according to the present interpretation of intention, we shall see whether this objection survives a synthesis of the three accounts. In case 4, my choice to give the five patients the coma patient's organs entails:

Double Effect: *"I intend" (I make a conscious effort to kill or let die—i.e., my primary aim is to kill in order to let live)*

Contrary Double Effect: *ditto*.

As we may begin to notice in these cases, Double Effect and Contrary Double Effect point to the formation of a common principle, what we may call the Synthetic Double-Effect Principle.

The Synthetic Double-Effect Principle: *Never aim or attempt to maximally violate agents' rights to freedom, and always aim at or attempt to not-maximally-violate agents' rights to freedom, even if it is in order to maximally uphold agents' rights to freedom.*

We might notice, in effect, that the maxim "Maximally violate an agent's right in order to maximally uphold an agent's right" entails a kind of contradiction that would not survive scrutiny via the PGC. Notice, too, that it does not follow from this that it is wrong to kill in self-defense. Why is this? This is because, and this is worth re-emphasizing, the focus here is on *aims*. The focus on aims entails an emphasis on intentions which, as we have previously analyzed in the case of self-defense, must conform to rules of logical consistency. One might recall that, in the

above-analyzed case of self-defense, I hypothetically asserted in killing my assailant, "I acknowledge my duty to uphold my freedom" and the lemma "I must uphold my right to freedom by doing only what is necessary to do so." My *aim* was to maximally uphold my right to freedom and not to maximally violate my assailant's right to freedom; the deliberative nature of my action was to *prevent* an aimed maximal violation of freedom, not to *initiate* one. The nature of such self-defense scenarios is also substantively different from cases 1, 2, 3, and 4 in that, once again, my set of possible choices is split between my approval of a maximal violation of the PGC and a duty-claim concerning my own freedom. In the aforementioned cases, by contrast, I do not face the dilemma of *accepting* or *rejecting* the intentions (or implicit assertions) of an agent whose survival is a necessary condition of some death. We can see from all of this, thus, how intensely intention-focused, and by consequence assertion-focused, Gewirth's deontology truly is. The PGC requires voluntary action in accordance with acceptable *assertions*. We must be careful, however, not to conflate intention and motive, as John Stuart Mill aptly pointed out.⁸⁴ Motive (in the Kantian sense) often does not coincide with intention and, in these cases, it is the intention that is significant to moral action. This distinction is important, for it could conceivably lead to such beliefs as that killing another person is obligatory if her motive for surviving some circumstance is merely to take revenge on another person. Right motives also matter, and we will see why when we further explore value theory. In the account of deontological ethics I have offered so far, however, motive is not directly relevant to right action. It is right intention, rather, that is directly relevant, as we have seen. I want to emphasize, however, that it is right intentions *concerning the consequences* that the PGC prescribes and that such an intentionalist approach does not at all preclude the idea of a concern for such consequences.

3.3.9 Conclusion and Synthesis

Now that we have investigated positions A, B, and C, which position should we accept? As it stands, Position C seems the most plausible, but we must consider the merits of the other two as well. In Position A, the argument of equal unjustifiability and, thus, equal *justifiability* of the choices in case 2 presented us with a kind of incompleteness or "moral freedom." In Position B, the Self-Contract in the presence of free agency was vital for understanding one reason why case 2 is relevantly different from case 1. Although Position A presents to us a claim of the equal unjustifiability (and thus the equal justifiability) of the large man's choices in case 2, the Self-Contract demonstrates why this might not be the case. In the spirit of Parfit, there would seem to be little in the way of justification for maximally acting against the interests of successive selves in order to prevent the maximal destruction

⁸⁴ John Stuart Mill, *On Liberty*, (Upper Saddle River: Longman, 2006).

of others' interests. Whether or not we accept Parfit's theory of personal identity, we can still accept this line of thinking, if only we consider the concept of "successive selves" to mean only our future self at various points in time. As with Parfit's arguments against Self-Interest Theory, which we surveyed in the section on negative justification, Position A instantiates a sort of unjustifiable particularism by asserting the arbitrariness of a certain kind of judgment in a certain kind of case. Position B showed this supposed arbitrariness to be unfounded. If it is unfounded, and the Self-Contract and free agency fill in this seeming gap of arbitrariness, then it seems there is some merit to Position B. This position on moral dilemmas is perfectly consistent with the PGC, but there is a serious question as to whether it is directly derivable from the PGC. As we shall see, any questions on such derivability will be resolved when we discuss virtue theory and the nature of goodness in Chap. 5. Position C, utilizing a somewhat Anscombean conception of intention directly derivable from the PGC, showed us that agency and levels of agency in these scenarios are relevant, but that it is intention that is even more relevant. Agency and levels of agency are relevant only to the extent that they demonstrate to us, in case 2, case 3, and relevantly similar cases, that we are not justified in pushing the large man over not simply because he has the power of choice and relevant rights to exercise his own choice, but because such rights cannot be overridden in this case precisely because the PGC implicitly prescribes that one attempt or try to save everyone while simultaneously attempting or trying to harm no one. The PGC is actually somewhat explicit: one must not aim to harm and one must aim not to harm, and one must also simultaneously aim to help. Obviously, this is not possible in such dilemmas, but the PGC is clear that one must attempt, to the greatest of one's abilities, to follow this directive. This, therefore, is the ultimate reason, on Gewirth's theory, why we should not push the large man over the bridge in case 2 and why we should not non-consensually take the organs from the person in case 3. Again, it is interesting that we find that these Gewirthian applications to moral dilemmas seem to confirm the spirit of Kant's original maxim concerning the impermissibility of treating agents as mere means to an end.

One notable critique of such a position on these types of moral dilemmas has come from Scanlon.⁸⁵ Philosophers, he claims, have thus far merely assumed the soundness of the Double-Effect Principle but have offered no proof to this effect. As such, we have no reason to believe that this principle, and by proxy the Contrary Double-Effect principle and the Synthetic Double Effect Principle, is sound. I believe that Scanlon is correct to state that philosophers have not actually proved the soundness of this principle, but he has overlooked some important points. Gewirth's account, as we have seen, corroborates both double effect and contrary double effect. Since Gewirth's theory, and as I will show Habermas's as well implicitly, prescribes the dual duty to attempt to help as many as possible while also attempting to harm none, their theories indirectly justify the principles of

⁸⁵ T. M. Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame*. (Cambridge: Harvard University Press. 2008).

double effect and contrary double effect. Gewirth's implicit emphasis on intentions as a justificans, via the justifiability of the assertions that derive from moral situations and which agents must implicitly either accept or reject, also demonstrates that intentions are directly relevant to moral justification. Scanlon, however, rejects just such a claim. He establishes a rather peculiar distinction between deliberative and critical uses of reasons and denies that persons can choose the reasons upon which they act. In the way Scanlon uses this distinction, I believe that he covertly conflates intention and motive, and thus fails to see exactly why intention, that is intention concerning the consequences, is directly relevant to the justification of an action. It is important, in the first place, to make a clear distinction between the goodness of right intention and the greater goodness of right motive. The effectuation of the best outcome (including by accident) makes an act by itself good, right intention makes an act better, while right motive makes a person good. I will claim in later sections and chapters that it is ultimately better to be a better person than merely to perform good acts. This will be supported by my later argument to the *summum bonum*. While he thus rightly states that intentions are irrelevant to whether outcomes are good, the judgment of intention and motive helps us judge whether outcomes are best, for included in the best outcome is the best intention coupled with the best motive. Thus, pushing the man off the bridge is bad (let us say, as an accident) because pushing the man off the bridge has the effect of sacrificing his occurrent and prospective freedom for those of others. This act, however, has *the same bad effect* as someone's accidentally flipping the switch in the simplified trolley dilemma. We must consider, however, that the mere deaths of the five people in both situations is also bad; nay, it is admittedly even worse than the accidental death of one person. The point here is that the performance of a bad or good *act* does not require any particular intention or motive at all. The performance of such acts, therefore, cannot be directly relevant to the goodness or badness of *behavior* as we think of it in normative terms. The goodness or badness of an *act* is not that which we are to judge in such situations; we must, rather, judge behavior. The judgment of behavior, however, requires the analysis of intention and motive. Scanlon cites an example of a physician who is about to perform an operation but who does not have the right "intentions" for doing so. On an intention-focused ethical view, he claims, we would have to accept that the physician should not do the operation and that the hospital should instead wait until it finds a physician with the right "intentions" in order to commence the surgery. This, once again, seems to indicate a conflation of intention and motive. If we replace Scanlon's mention of intention here with that of motive, his critique makes complete sense. We do not, however, want to state that it does not matter whether or not the physician has the right intention, for, on my account, a judgment concerning the justifiability of the act in question, for whatever underlying reason, is what constitutes one's intention. How, then, can we even have an ethics without intention's being the central feature? Without intention, on my account disconnected from underlying reasons, voluntary actions could not even be undertaken because there would be no rationales for undertaking them. Certainly, we could still judge the goodness and badness of consequences and actions, but we would not be

judging their morality and immorality. The concept of morality only makes sense as constitutive of a social framework within which agents make decisions. Thus, it is *decisions* that are ultimately morally relevant and not the acts themselves. On this account, then, we can see that agent A's pushing the man off the bridge is impermissible (which is a moral concept) because doing so would constitute, regardless of agent A's reason for doing so, using this man "merely as a means," in that agent A's pushing him off the bridge would entail the truth of the statement, "It is permissible to extinguish this man's agency as a means to preserving the agency of others," which in fact is untrue. It is the falsehood of the statement entailed by the *action*, and not by the *agent*, that makes this act impermissible. Thus it is the *nature* of the act that it just *is* a use of an agent in this way by agent A which makes the act impermissible. We must remember, however, that the *nature* of the act involves the *nature* of the voluntary action undertaken by agent A. In other words, a central part of the morality of the act consists in the nature of the statements entailed by the activity of the agent. If, then, I both do everything I can to effect benefit (saving the five) while simultaneously doing everything I can do avoid doing harm (not pushing the man over), then I have acted morally in not pushing the man over. It is key here to emphasize that *I*, judged on my merits *as an agent*, have acted *morally*. It is an altogether different thing to say that a good act merely occurred. These agentive factors are that in which intention, and not motive, consists. My actions exhibit and are reflective of my intentions; this is precisely one of Anscombe's theses. If, in the bridge case, the five people have died, it was not because I failed to act to save them but because I could not save them without *attempting* to, *trying* to, or *aiming* at killing the man by pushing him off the bridge. If, in the simplified trolley dilemma, the one person died, it was not because I *attempted* to, *tried* to, or *aimed* at killing him but because I *attempted* to, *tried* to, and *aimed* at saving as many people as possible while at the same time *attempting* to, *trying* to, and *aiming* at harming *no one*.⁸⁶ This pattern of intention is also completely consistent between both such cases. This is what I believe Scanlon overlooks. Thus, we see that an intention is the normative statement or reason entailed by a proposed course of action, while a *rationale* is that justification which the agent actually proposes for the action. Rationales, Scanlon could then rightly claim, are not directly relevant to whether the agent should actually do the action. The agent should do the obligatory action because a certain normative statement concerning the obligatory nature of the proposed action is entailed by the situation; the rationale is not.

These classic moral dilemmas are instructive, for as we shall see, the results we have derived here from Gewirth's theory are the very same conclusions we would reach in Habermas's theory of discourse ethics. In the following analysis, we shall see that Habermas's theory is, interestingly, a sort of mirror image of Gewirth's.

⁸⁶ It is key here that I am not aiming at harming merely as few people as possible; this is simple consequentialism. Instead, I am aiming at saving as many people as possible without harming anyone. This is implicit in the norms derived from the PGC.

While Gewirth derives his conclusions from the perspective of the subject and her or his self-contradiction, Habermas starts from the opposite end and derives his conclusions from inter-subjective discourse and alter-contradiction. As we shall see, both theories, when taken together, provide a powerful account of the underlying logical structure of right action.

As we shall also see, Gewirth and Habermas's theories address intention, whereas I will ultimately address motive. On Gewirth and Habermas's theories, right intention is necessary for right and good action. Scanlon's analysis covertly only addresses right motives, but, as I have argued along with Scanlon, right motives are irrelevant to right action. We must remember that what "intention" implies within the context of Gewirth and Habermas's respective theories is more comprehensive than in other theories. It implies, for the agent, assaying the situation, analyzing the various assertions entailed by each possible course of action, identifying the true assertion in accordance with the PGC, or in Habermas's case in accordance with principles (D) and (U), and then acting in accordance with the dictates of that true assertion. Again, however, it is not necessary for the agent's fulfillment of a right action that he actually make his decision based upon whether it is in accord with the PGC; it is sufficient merely that his act be consistent with the PGC or, in Habermas's case, with the principles that constitute the Ideal Speech Situation. This will be important to remember as we progress in this cumulative analysis.

3.4 A Sound Positive Account, Part II: An Analysis of Habermas's Discourse Ethics

In this section, I analyze Habermas's theory of communicative ethics. I demonstrate that, despite his intentions, Habermas's theory is ultimately both foundationalist and monological. As such, I identify it as employing much the same method as Gewirth's theory; it is, in fact, a mirror image of Gewirth's theory. It too, I explain, is both sound and substantive. It, however, bears the same flaws as Gewirth's theory in that it does not provide a response to what I have called general questions of value. Habermas's theory, however, likewise reveals fundamental truths about the nature of value judgment and practical rationality. Combining these Gewirthian and Habermasian insights on value judgment yet rejecting their theoretical foundations, I proceed toward new foundations.

Jürgen Habermas is in some ways an enigmatic and perplexing thinker. This is not, however, due to the unclarity or obscurity of his prose, the incoherence of his concepts and distinctions, or the complete incompatibility of critical theory with other strains of thought. On the contrary, when one reads Habermas, one is struck by the relative clarity of the text and the coherence of the arguments. One cannot always say the same for Habermas's Frankfurt School predecessors. In light of his well-known synthesis of critical theory and analytic philosophy, it is also vital to

emphasize that these philosophies are not completely incompatible. While he adopts the critical-theoretical view concerning the inter-subjectivity of truth, Habermas coherently uses the tools of analytic philosophy to better understand the lifeworld. This synthesis of critical theory and such foundationalist paradigms as analytic philosophy and Kantianism, however, ultimately conflicts with his anti-foundationalist project and, in an intriguing way, infects his thought. While, in his theory of communicative action more generally, he is able to accommodate various elements of foundationalist philosophies while preserving the anti-foundationalist premises of critical theory concerning the inter-subjectivity of truth, Habermas's normative theory cannot remain unproblematic while preserving this critical-theoretical conception of truth. While he attempts to import this conception of truth into his normative framework, and while he often does so convincingly, Habermas's normative theory is ultimately covertly not only foundationalist but also monological. Habermas attempts to straddle critical theory, analytic philosophy, and Kantianism while preserving key normative elements of each. His interpretation of the critical-foundationalist juncture with regards to normativity, however, contains irresolvable tensions. These tensions resulting from the conflict between his foundationalist and monological tendencies and his critical-theoretical claims, however, are ultimately his strength and not his weakness. As we shall see, it is actually Habermas's anti-foundationalist, critical-theoretical approach which, ironically, substantively grounds the powerful idea of epistemological uncertainty that is central to an ethics of anti-paternalism, non-aggression, and liberation and which, as Beyleveld has demonstrated through his Precautionary Principle, is also central to Gewirth's theory.

Habermas's critical theory owes a large debt to the Hegelian tradition. As thinker Rüdiger Bubner points out, it is Habermas's crucial distinction between work and interaction, as well as his subsequent emphasis on interaction, which singles out his work as characteristically Hegelian.⁸⁷ Interaction, conceptualized along the lines of the dialectic, is the *telos* of language and speech for Habermas, and, like the dialectic, emancipation is possible only through interaction leading to "universal equality."⁸⁸ Habermas, however, goes even further than Hegel in some respects with regards to the concept of interaction. Whereas Hegel views work as a crucial liberatory force alongside interaction, Habermas emphasizes, throughout his work, the primacy of interaction and its importance even to work itself.⁸⁹ Thus, far from a half-hearted critical theorist, Habermas represents one of the most earnest adherents to the critical-theoretical notions of historicism and cultural contextualism, as well as to the discursive nature of truth that proceeds from these key bases of social

⁸⁷ Rüdiger Bubner, "Habermas's Concept of Critical Theory." *Habermas: Critical Debates*. Ed. John B. Thompson and David Held. (Cambridge: The MIT Press. 1982).

⁸⁸ *Ibid.*, 49.

⁸⁹ Seyla Benhabib, *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory*. (New York: Columbia University Press. 1986); Robert C. Holub, *Jürgen Habermas: Critic in the Public Sphere*. (London: Routledge. 1991); Thomas McCarthy, *The Critical Theory of Jürgen Habermas*. (Cambridge: The MIT Press. 1978).

critique. It is this central idea of a discursive, non-objective truth derived from an emphasis on interaction, however, which is fundamentally at odds with Habermas's attempt to incorporate strains of foundationalist thought, including Kantian and analytic philosophy, into his normative work.

This internal tension is most apparent in Habermas's meta-ethical and ethical theory. In *Moral Consciousness and Communicative Action*, he establishes an ethical theory centered on what he calls "moral validity."⁹⁰ This is expressly different from moral or normative truth. Since truth is discursive, and thus contextual and historicist, there can be no normative truth, as it were. Instead, there can be only valid normative claims. While superficially similar to Hare's normative claims, the differences between Hare's ethics and that of Habermas are crucial. While Hare claims that positive statements of value are merely statements of personal preference or "commendation" and thus must be evaluated solely on the basis of their formal validity, Habermas presents a far more substantive account of such judgment. Moral judgments, for Habermas, are claims to validity on the part of participants in a discourse. "To say that *I ought* to do something," he states, "means that I *have good reasons* for doing it."⁹¹ These good reasons, however, are not definitively valid unless all others affected by the action in question can, as equal participants in a discourse, validate these reasons. For Habermas, therefore, morality is inherently discursive and dialogical. This notion is reinforced in Habermas's philosophy by his use of Kohlberg's theory of moral development, which serves to strengthen Habermas's claim that notions of morality are psychologically founded, from infancy, upon interaction and thus upon discursive validity. This idea of discursive validity, however, comes dangerously close to undermining the basic assumptions of critical theory concerning anti-foundationalism, historicism, and contextualism. While it is clear that the notion of discursive validity conforms to the critical-theoretical interpretation of truth as historically and socially situated, Habermas's highly formalistic, process-oriented, conception of ethics establishes an underlying logical structure that is inherently both monological and suggestive of claims to trans-historical, trans-cultural, objective soundness. While in the context of critical theory, Habermas's claim to inter-subjective normative validity undermines itself and in fact suggests objective normative soundness, within the context of analytic philosophy, this self-defeating nature is wholly unproblematic. Habermas, therefore, is precariously situated between critical theory and foundationalist philosophies, including those of Kant and the analytic philosophers.

Within the context of European intellectual history, Habermas's predicament is significant. His ethical thought suggests a strong motivation to keep ethics as fragmented, and thus as fluid and dynamic, as possible. The critical-theoretical impulses toward subjectivity and inter-subjectivity most evidently appeal to this motivation. The explanations for this motivation can be at least partially accounted for by considering Habermas's post-WWII involvement in West German politics and political thought, which brought him face to face with the oppressive potential of

⁹⁰ Habermas, *MCCA*, (Cambridge: The MIT Press. 1990): 49.

⁹¹ *Ibid.*

rigid, foundationalist claims to ultimate, objective value. At the same time, however, Habermas's increasing incorporation of elements of Kantian and analytic philosophy into his critical-theoretical stance over time might suggest a realization on his part that a complete epistemological and normative escape from the foundational and the monological is ultimately both impossible and impracticable. Although this is undoubtedly an incomplete explanation for the critical-foundationalist tensions we see in Habermas's thought, for they do not directly account for the agency involved in his intellectual decisions, these explanations may provide a starting point for the historical and philosophical reinterpretation of Habermasian ethics.

3.4.1 *Critical Theory, Interaction, and Communicative Ethics*

Numerous historians and philosophers have assayed Habermas's ethical thought and its historical roots. Not surprisingly, the majority of these are intellectuals who largely sympathize with the critical-theoretical project.⁹² Few intellectuals of the analytic tradition, save those directly involved in debates with Habermas such as Rawls and Gewirth, have given Habermas serious consideration. These intellectuals are certainly at a loss, for Habermas's work has been instrumental to the re-conceptualization of critical theory toward a linguistically oriented form that places far more consideration than did his predecessors upon speech-act theory and universal pragmatics.⁹³ This integration of analytic-philosophical elements has been coupled with an equally important move on Habermas's part, namely the transition from a work-oriented conception of liberation and emancipation to an interaction-oriented one. Seyla Benhabib, a student of Habermas, emphasizes the magnitude of the break Habermas has made with earlier critical theorists such as Max Horkheimer, Theodor Adorno, and Herbert Marcuse. Because these philosophers saw rationalization as an "all-encompassing process in which the organization of the productive forces, societal institutions, cultural meaning patterns, and personality structures all submit to the same logic of ever-increasing fragmentation, atomization, efficiency, and formalism,"

⁹² Benhabib, *Critique, Norm, and Utopia*; Bubner, "Habermas's Concept of Critical Theory"; Holub, *Jürgen Habermas*; Martin Jay, *The Dialectical Imagination: A History of the Frankfurt School and the Institute of Social Research, 1923–1950*. (Berkeley: University of California Press. 1973); McCarthy, *The Critical Theory of Jürgen Habermas*; William Rehg, *Insight and Solidarity: A Study in the Discourse Ethics of Jürgen Habermas*. (Berkeley: University of California Press. 1994); Matthew G. Specter, "Habermas's Political Thought, 1984–1996: A Historical Interpretation." *Modern Intellectual History*, 6, no.1 (2009): 91–119; *Habermas: An Intellectual Biography*. (Cambridge: Cambridge University Press. 2010).

⁹³ Jürgen Habermas, "Hermeneutic and Analytic Philosophy. Two Complementary Versions of the Linguistic Turn?" *German Philosophy Since Kant*. Ed. Anthony O'Hear. (Cambridge: Cambridge University Press. 1999); Holub, *Jürgen Habermas*; McCarthy, *The Critical Theory of Jürgen Habermas*; Specter, *Habermas: An Intellectual Biography*.

they viewed rationalization, and by proxy rationality, in a monolithic, and thus a monolithically destructive, way.⁹⁴ The earlier critical theorists, she suggests, sought to conceptualize rationality in a holistic manner and often ended up conceptualizing it in a monochromatic manner instead. Habermas himself makes this point in *The Philosophical Discourse of Modernity*.⁹⁵ Habermas, by contrast, makes a clear distinction between what he calls "system integration" and "social integration." "By the former," Benhabib explains, "is meant a connection of action consequences, whereas 'social integration' refers to the coordination of action through the harmonizing of action orientations."⁹⁶ This is far from the "one-dimensionality" that Marcuse proposes as the characterization of the relationship between reason and society and the self-destruction that reason represents in Horkheimer and Adorno's philosophies.⁹⁷ It is this revitalized conception of reason (*Vernunft*) that Habermas uses to found his idea of interaction. The *telos* of interaction envisioned in Habermas's ethics is the universal equality envisioned in that of Hegel, coupled with a reinforced conception of equality in terms of power relations. This critical-theoretical concept of interaction, and thus his reconceptualization of reason, is the basis of his ethical thought.

In Habermas's ethics, a norm is only justified if it can endure the rational scrutiny of all those affected by the norm. As Habermas states: "Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*."⁹⁸ This is what Habermas calls the Principle of Discourse (D). His second principle which follows from his formulation of moral validity is the Principle of Universalization (U): "All affected can accept the consequences and the side effects its *general* observance can be anticipated to have for the satisfaction of *everyone's* interests (and these consequences are preferred to those of known alternative possibilities for regulation)."⁹⁹ It is when these conditions are met that all participants in discourse are in what Habermas calls the Ideal Speech Situation. This situation, however, is more than a mere description of a possible state of interaction. It is, rather, a state with strong normative legitimacy. "If these considerations are to amount to more than a definition favoring an ideal form of communication and thus prejudging everything else," he argues, "we must show that these rules of discourse are not mere *conventions*; rather, they are inescapable presuppositions."¹⁰⁰ It is when Habermas delves more deeply into the justification for the Ideal Speech Situation, however, especially via this idea of "inescapable presuppositions," that his theory becomes increasingly problematic within the context of critical theory.

⁹⁴ Benhabib, *Critique, Norm, and Utopia*, 230.

⁹⁵ Jürgen Habermas, *The Philosophical Discourse of Modernity: Twelve Lectures*, (Cambridge: The MIT Press. 1990).

⁹⁶ Benhabib, *Critique, Norm, and Utopia*, 230.

⁹⁷ Max Horkheimer and Theodor W. Adorno, *Dialectic of Enlightenment: Philosophical Fragments*. (Stanford: Stanford University Press. 2002).

⁹⁸ Habermas, *MCCA*, 66.

⁹⁹ *Ibid.*, 65.

¹⁰⁰ *Ibid.*, 89.

Philosopher William Rehg emphasizes the basic idea of critique, central to critical theory, that is also inherent in Habermas's ethics. "The strong discourse-ethical notion of autonomy," he argues, "requires subjects to question even pregiven legitimating frameworks and authorities. Validity can then be redeemed only on the basis of the formal properties of argumentation."¹⁰¹ Once again, it is the normative force of Habermas's "inescapable presuppositions" which legitimates this discourse-ethical autonomy. This autonomy or freedom, Rehg reveals, is profoundly prescriptive not only in interaction or dialogue, but also in monologue. "Such freedom," he states, "would seemingly have to include not only a freedom from external coercion or manipulation but from internal compulsions and self-deceptions as well."¹⁰² Rehg strongly suggests that the "inescapable presuppositions" of Habermas's theory refer to a highly formalistic process. Indeed, intellectuals have long singled out Habermas's entire conception of critical theory as one that is extremely process-focused and relatively content-light.¹⁰³ These "inescapable presuppositions," then, are unmistakably supported by an underlying formalistic logic. Rehg emphasizes this with regards to what he calls Habermas's "No Alternatives Thesis," according to which "the metavalue of autonomous cooperation is without viable alternatives."¹⁰⁴ Rehg makes this argument in response to critics of Habermas's theory who object that Habermas makes an unjustified claim concerning the value of the Ideal Speech Situation itself and, thus, that his argument is circular. It is because of these criticisms, then, that Rehg emphasizes the formalistic nature of Habermasian ethics and, subsequently, that the Ideal Speech Situation is "a unique sort of metavalue that one cannot rationally reject except in an abstract fashion."¹⁰⁵ He tests two kinds of theories against Habermasian ethics in order to demonstrate a deeper point concerning the underlying formalistic logic of Habermas's theory. The first, Self-Interest Theory, which he labels (SI), and the second, a group value held by an agent solely because the group holds that value, which he labels (GV), bring to light the core nature of Habermas's "inescapable presuppositions." If agents hold SI or GV in opposition to other agents yet refuse to engage in discourse to reach an understanding, then, Rehg states, "agents. . . must therefore deal with the fact that other persons, to all appearances just as rational, object to their chosen course of action. There thus exists a contradiction between the reflective attitude demanded by moral questions and an attitude consistently oriented by (SI) or (GV)." In this light, Rehg asks, "How can agents following (SI) or (GV) deal with the rationality of others?"¹⁰⁶ In *Moral Consciousness and Communicative Action*, Habermas refers to such a refusal to

¹⁰¹ Rehg, *Insight and Solidarity*, 35.

¹⁰² *Ibid.*, 43.

¹⁰³ Benhabib, *Critique, Norm, and Utopia*; Bubner, "Habermas's Concept of Critical Theory"; Holub, *Jürgen Habermas*; McCarthy, *The Critical Theory of Jürgen Habermas*; Rehg, *Insight and Solidarity*; Specter, *Habermas: An Intellectual Biography*.

¹⁰⁴ Rehg, *Insight and Solidarity*, 140.

¹⁰⁵ *Ibid.*, 147.

¹⁰⁶ *Ibid.*, 166.

participate in discourse as a “performative contradiction,” for one’s refusal to take the opinions into account of an agent whose interests are affected by one’s decision is tantamount to a denial of that agent’s agency.¹⁰⁷ There are, of course, counterarguments to this, and we will explore them in due course.

Habermas is careful, Benhabib explains, not to claim that performative contradiction is equivalent to literal or logical contradiction. Habermas’s argument, in other words, is not aprioristic or deductive. One of Habermas’s main admitted aims is to reconstruct Kantian ethics, along with its principle of universalization, in light of Hegel’s criticisms of Kant and the subsequent tradition of critical theory while still retaining the legitimacy of reason in the practice of philosophy.¹⁰⁸ It has proven difficult for Habermas to straddle both positions while maintaining the integrity of both. Benhabib’s critique of Habermas addresses some of the issues concerning his attempt to attain these simultaneous goals. Although not explicit, Benhabib’s critique addresses Habermas’s underlying Kantian impulse. “Prior to the non-foundationalist turn of the *Theory of Communicative Action*,” she reveals, “it was argued that there was a less than contingent, quasi-transcendental connection between the structures of rational speech and a communicative ethic. This strong justification program, which would *ground* a communicative ethic on the fundamental norms of rational speech, was untenable, and in his later formulations Habermas has moved away from it.”¹⁰⁹ Benhabib goes on to explain that Habermas’s original conception of communicative ethics was, in fact, a wholly deductive one in which “the conditions of an Ideal Speech Situation *entailed* the acceptance of certain ethical norms.”¹¹⁰ At a certain point, she argues, it becomes difficult for Habermas to make the claim to transcendentalism or even quasi-transcendentalism. The original intention of his theory, it is clear, was aprioristic and foundationalist. It is, therefore, highly questionable, even in light of his genuine attempt at historicism and contextualism, whether Habermas is able to remain a critical theorist while pursuing a neo-Kantian or analytic ethics.

The last hope for the possibility of a Habermasian hybrid between critical-theoretical ethics and strongly Kantian and analytic components of ethics, it then seems, lies in a vital distinction through which Habermas implicitly links his theoretical framework with Hegel’s. That distinction, between the right and the good, is strikingly similar to Hegel’s. The distinction is furthermore pronounced since, in the history of Western ethics, such a distinction between the right and the good has seemed superfluous. Habermas quite firmly makes the distinction between ethics (*Sittlichkeit*) and morality (*Moralität*), and the accompanying concepts of the good and the right, respectively.¹¹¹ Hegel characterizes ethics as “the idea of

¹⁰⁷ Habermas, *MCCA*, 87.

¹⁰⁸ *Ibid.*, 5, 8, 63.

¹⁰⁹ Benhabib, *Critique, Norm, and Utopia*, 287.

¹¹⁰ *Ibid.*, 294.

¹¹¹ Habermas, *MCCA*; Holub, *Jürgen Habermas*; McCarthy, *The Critical Theory of Jürgen Habermas*; Rehg, *Insight and Solidarity*; Specter, *Habermas: An Intellectual Biography*.

freedom.”¹¹² The idea of freedom embodies ethics, in Hegel’s view, through the idea of the good. As he states: “In ethical life as a whole, both objective and subjective moments are present, but these are merely its forms. Its substance is the good, that is, fulfilment of the objective [united] with subjectivity.”¹¹³ For Habermas, this is the realm of value, or as he seems to characterize it both in *Moral Consciousness and Communicative Action* and in his objections to Rawls’s political theory, as the realm of mere preference.¹¹⁴ This component of Habermas’s characterization of the good, admittedly, differs much from Hegel’s in that Hegel views the good not as mere preference but as “*realized freedom, the absolute and ultimate end of the world.*”¹¹⁵ The key similarity, however, between Habermas and Hegel as regards the concept of the good is their shared view of the good as objectivity united with subjectivity and concerned with the welfare of the subject. Hegel’s conception of morality and right are also extremely similar to Habermas’s. For Hegel, morality and right are distinguished from each other in their positing of obligatory prescriptions and proscriptions, respectively. Morality, for Hegel, concerns “the determination whereby the individual will within it is superseded; and in consequence, since the determination of one-sidedness disappears, two wills with a positive reference to one another are now posited.”¹¹⁶ Again, Habermas and Hegel differ somewhat on their conceptions of morality and right, but both philosophers accept certain core concepts inherent in both. While Habermas holds that questions of right and justice belong to the sphere of morality, both hold that morality is the sphere in which questions concerning obligations to collective welfare or general interest are operative.

For Habermas, this adoption and reinterpretation of Hegel’s distinction aids him in adopting certain Kantian and analytic positions while continuing to hold on to some semblance of critical theory via a claim to anti-foundationalism. As we have already seen, much of the criticism of Habermas’s theory on the part of philosophers derives from the claim that, in arguing for the normative legitimacy of the Ideal Speech Situation, Habermas makes an unjustified assumption concerning the normative value and, thus, the obligatory nature of the Ideal Speech Situation. It is to Habermas’s advantage, then, to make the ethics-morality and good-right distinctions, for he can claim from this position that he is not making an assumption concerning the value or goodness of the Ideal Speech Situation because “value” and “goodness” refer to the particularist preferences of the individual subject. The moral, the right, and the just, however, are intrinsically embodied within the Ideal

¹¹² G. W. F. Hegel, *Elements of the Philosophy of Right*. (Cambridge: Cambridge University Press, 1991): 189.

¹¹³ *Ibid.*

¹¹⁴ Habermas, *MCCA*; “Reconciliation Through the Public Use of Reason: Remarks on John Rawls’s Political Liberalism,” *Habermas and Rawls: Disputing the Political*. Ed. James Gordon Finlayson and Fabian Freyenhagen. (New York: Routledge, 2011).

¹¹⁵ Hegel, *Elements*, 157.

¹¹⁶ Hegel, *Elements*, 140.

Speech Situation. From this perspective, therefore, Habermas can characterize obligation without referring to foundationalist-sounding, “inescapable presuppositions.” Indeed, even in his political theory, he uses such deontic concepts as the right and the just in order to guide and delineate political behavior and not to provide an ultimate foundation for it.¹¹⁷ Such deontic concepts, instead, merely embody the ethical interactions to which legitimate politics must adhere. This approach, however, still does not allow Habermas to escape the “inescapable presuppositions” which threaten his simultaneous adherence to critical theory, analytic philosophy, and Kantianism.

3.4.2 *Habermas, Gewirth, and the Inescapability from the Monological*

Despite his best efforts at concealing his foundationalist presuppositions, such ideas as the embodiment of justice and right in the discursive procedure only reinforce the suspicion of a foundationalist tension in Habermas's theory. One can still ask the question, for instance, why one ought to accept Habermas's conception of obligation. If normative truth, or normative validity, is truly discursive and context-dependent, then there seems to be no reason why an egoistic or a might-makes-right conception of obligation would not be equally as acceptable as an altruistic one. Habermas, however, can respond to this claim the same way he addresses similar claims in *Moral Consciousness and Communicative Action*. These conceptions of obligation are wholly unacceptable for Habermas because, as he explains, “the justification of norms and commands requires that a real discourse be carried out and thus cannot occur in a strictly monological form.”¹¹⁸ In other words, normative justification simply *is* discursive justification. One cannot, in other words, make the validity claim “I ought to do X,” even monologically, without presupposing the normative legitimacy of the discursive lifeworld and the entitlements of other rational beings to criticize this claim. As Habermas states: “Every person who accepts the universal and necessary communicative presuppositions of argumentative speech and who knows what it means to justify a norm of action implicitly presupposes as valid the principle of universalization, whether in the form I gave it above or in an equivalent form.”¹¹⁹ The aforementioned is thus Habermas's reformulation of Kant's principle of universalization in action. If this simply *is* the nature of justice and right, however, then it seems clear that this amounts to a strong trans-cultural, trans-historical claim. If this simply *is* the nature

¹¹⁷ James Gordon Finlayson and Fabian Freyenhagen, “Introduction,” *Habermas and Rawls: Disputing the Political*. Ed. James Gordon Finlayson and Fabian Freyenhagen. (New York: Routledge. 2011): 9.

¹¹⁸ Habermas, *MCCA*, 68.

¹¹⁹ *Ibid.*, 86.

of justice and right, and if such conceptions of obligation as that of a Thrasymachus, Thucydides, Callicles, or Machiavelli are illegitimate, then they have always been illegitimate, will always be illegitimate, and are illegitimate everywhere. Underlying Habermas's supposedly critical-theoretical account of ethical theory, then, are covert norms which make a claim to objective truth. In fact, Habermas's argument is fundamentally aprioristic and deductive. The idea of performative contradiction, upon closer analysis of Habermas's highly formalistic approach, becomes mere logical contradiction. If we simplify Habermas's argument to its most elementary claims, we see that the argument is essentially that, if agent A makes a validity claim whose fulfillment would affect agent B, then if agent B disagrees with agent A's validity claim, then since all agents are rational beings who are thus capable of demonstrating an invalid validity claim to be invalid, agent A can only act on her validity claim by accepting the antecedent claim that agent B is not an agent. This formulation involves more than merely performative contradiction; it involves making literal ought-judgments which are *a priori* false via logical contradiction. As Benhabib points out: "Such reconstructions continue the task of transcendental philosophy with altered means. The lines separating so-called 'a priori' conditions from a posteriori rules of competent behavior get increasingly blurred, for the justification for calling some condition or feature of our action and speech capacities 'a priori' is simply that it corresponds to the implicit know-how of competent actors, for which we see no possible alternatives at the present."¹²⁰ Although Benhabib is wary of Habermas's approach, that his theory conflicts with the conception of truth set forth by critical theorists is no objection to this theory as a whole. Instead, it is merely further evidence that Habermas fails to accommodate simultaneously the critical-theoretical, analytic-philosophical, and Kantian positions.

Even more significantly, Habermas's theory is not even fully dialogical, as he would like. If he wishes to maintain the claim that his conceptions of obligation, justice, and right are embodied within the Ideal Speech Situation and, thus, that the Ideal Speech Situation is self-justifying, Habermas has no choice but to defend a monological position. Although it is one of Habermas's principal claims that, in order to justify a norm, an agent must validate her norm against the input of others whom the norm affects, he seems to ignore the fact that, on his own account, it follows from this that all agents who engage in communicative action must logically accept this formulation. In other words, in order for Habermas's dialogical formulation to be valid, he must likewise validate it against the input of other agents and have them, from the perspective of their own subjectivity, accept the formulation as valid. Thus, paradoxically, in order for the dialogical formulation to be valid, a monological formulation about the dialogical formulation must also be valid from the perspective of other agents whom Habermas wishes to convince. More specifically, all agents must accept the aforementioned formulation: "If agent A makes a validity claim whose fulfillment would affect agent B, then if agent B disagrees with agent A's validity claim, then since all agents are rational beings who are thus

¹²⁰ Benhabib, *Critique, Norm, and Utopia*, 295.

capable of demonstrating an invalid validity claim to be invalid, agent A can only act on her validity claim by accepting the antecedent claim that agent B is not an agent." Thus, from the perspective of all other agents, if they do not accept that all other agents are entitled to an equal amount of autonomy in discourse, they effectively claim that those agents are not agents and thus contradict that those agents are agents.

A comparison with Gewirth's ethical theory demonstrates the astounding parallels between his and Habermas's claims. Whereas Habermas claims a dialogical position and a discursive conception of normative truth, Gewirth is quite open about his own position's monological, deductive, and aprioristic pretensions. We might recall that Gewirth calls his own approach a "dialectically necessary method" in order to emphasize that his argument is not assertoric, that is, that it does not make direct claims to objectivity but merely analyzes the propositions which follow from the agent's subjectivity. As such, his approach seems somewhat to mirror the approach which Habermas takes with regards to subjectivity and inter-subjectivity. This is because Gewirth's approach, like that of Habermas, is similarly decentered with regards to objectivity and the assertoric. At the same time, however, Gewirth ultimately wants to make claims to normative objectivity.

Once again, on Gewirth's view, voluntariness and purposiveness, as the constitutive features of agency, logically entail the acceptance of certain assertions on the part of all agents. He begins with the most basic assertion: "I do X for purpose E." Since all agents also make normative claims concerning their actions, even if they are merely instrumental ones, such claims as "I ought to do X" and "I must do X" follow from this assertion. Since it is also true that, in order to engage in such action, all agents must also make normative claims to the freedom to do those things, it follows that agents must judge "I must be free to do X." Thus, all agents implicitly make the claim "I must have freedom." From this deontic claim, Gewirth argues, follows the dialectically necessary claim "I have a right to freedom." Interestingly, up to this point, Habermas seems to have little issue with Gewirth's account. It is only when Gewirth makes the deductive move from the necessity claim to the right claim that Habermas interjects. "Alan Gewirth," he states, "tries to derive basic ethical norms from the structures and general pragmatic presuppositions of goal-directed action. He applies presuppositional analysis to the concept of the capacity to act spontaneously and teleologically, in order to show that every rational actor is forced to view as goods the latitude he has to act, or more generally, all resources for the realization of his ends. Interestingly, however, the teleological concept of action is inadequate to provide a transcendental-pragmatic justification of the notion of a *right* to such 'necessary goods,' as opposed to the idea of the goods themselves."¹²¹ While Habermas rejects this logical move, it is most salient here that Habermas seems actually to accept nonetheless Gewirth's justificatory procedure more generally. As he admits: "If instead the concept of communicative action is chosen as a basis, one gets, via the same methodological route, a concept

¹²¹ Habermas, *MCCA*, 101.

of rationality strong enough to *extend* the transcendental-pragmatic derivation of the moral principle to the basis of validity of action oriented toward reaching an understanding.”¹²² Gewirth, much like Habermas, completes his argument by introducing the principle of universalization. Thus, he reasons, if all agents must make these normative claims, then it follows that they must make the same normative claims with regards to all other agents on pain of contradiction. Once again, Gewirth is explicit: if an agent contradicts herself in this regard by claiming that other agents do not have the generic rights, then her claim is not simply invalid; it is *a priori* false. It is difficult, then, from this perspective, to demonstrate that Habermas does not make effectively the same sorts of claims. On the contrary, it is clear that the foundationalist impulses of analytic philosophy and of Kant usurp critical theory within his normative model.

3.4.3 *Clarifying the Mirror Image*

We may make what I have called “the mirror image” that constitutes Gewirth’s and Habermas’s respective theories a bit clearer by dealing with an objection to Rehg’s and my own interpretation of Habermas’s theory. One may recall my earlier exposition of Rehg’s account of Habermas’s theory. There, we surveyed the claim that, “If agent A makes a validity claim whose fulfillment would affect agent B, then if agent B disagrees with agent A’s validity claim, then since all agents are rational beings who are thus capable of demonstrating an invalid validity claim to be invalid, agent A can only act on her validity claim by accepting the antecedent claim that agent B is not an agent.” It may be objected that one can fail to recognize an agent as an agent without, in principle, contradicting that that agent is an agent, for it might be claimed that one cannot contradict that which one cannot in principle know, or at least that which one is not entirely certain of, as one is certain of one’s own agency. This is a very interesting and in fact important objection, and it is one which brings us squarely to the kind of analysis that is vital for the theoretical unification of Habermasian and Gewirthian theory.

Beyleveld himself has expressed much doubt as to whether a contradiction is instantiated by the failure to act in accordance with the Ideal Speech Situation. “It is . . . possible,” he states, “to understand discourse ethics . . . as claiming that the rules/principles of discourse ethics are rules that anyone who claims to engage in a discourse with another must accept on pain of contradicting the idea of being engaged in a discourse. . . . However, such an argument, even if valid, would lack the full dialectical necessity Gewirth claims for the PGC, as there would be no contradiction, without more being established, in an agent denying the existence of

¹²² Ibid.

other agents. Nevertheless, such an argument would be very rationally compelling."¹²³ While Habermas and Apel have offered substantive counterarguments concerning the impossibility of ethical solipsism due to the implicit transcendental-pragmatic acknowledgment of the lifeworld and, thus, of the existence of the other, it is indeed possible to travel unnecessarily far down the rabbit hole of metaphysics in attempting to derive answers to normative questions. With regards to solipsism, suffice it to say that I believe Habermas and Apel bear the stronger argument here, for it is non-nonsensical to speak, at the very least, of *moral* or *ethical* theory without implying an attempt to answer questions concerning alterity. For instance, even in posing the central moral question of egoism, namely "Shall I give moral consideration to myself only or to others in addition to myself?," one implicitly presupposes and assumes the existence of others. Thus, whether or not it can be shown metaphysically that there exist other agents (which in any case I believe it probably can be so shown), it suffices to make the assumption, as in a formal proof by cases, that other agents exist. Thus, we may ask the question, "Given that other agents exist, how should I treat them?," and I believe we may offer substantive responses to this question without proving, beyond metaphysical doubt, that other agents actually exist.

The objection previously posed, however, makes a far less radical assumption than solipsism. It instead leaves room for the existence of other agents but offers the claim that it is in principle impossible for me to know with full certainty that an entity which I encounter is indeed an agent. Recall that we previously surveyed such a problem in examining Beyleveld's interpretation of the Principle of Proportionality. There, we saw that his interpretation was derived from his Principle of Precaution, which is founded upon the realization that it is empirically impossible to determine with full certainty whether a given entity is or is not an agent. It is in Beyleveld's Principle of Precaution, I claim, that we find a partial link between Gewirthian theory and Habermasian theory and begin to refine and clarify the "mirror image" I have been claiming is constitutive of these two models. The Principle of Precaution, Beyleveld claims, derives directly from Gewirth's theory in its elaboration of what the PGC dictates must be done in cases of interpersonal conflict. In fact, he states, such a principle yields the justification of a "procedural morality" and subsequently procedural structures for determining how disputes over the validity of varying beliefs and value systems ought to be settled, structures and procedures of which have been agreed upon consensually by all concerned beforehand.¹²⁴ As we begin to see, this is remarkably similar to the conclusions which are derived from Habermasian theory and, of course, more explicitly from Apel's discourse ethics. While Apel's theory is important to discourse ethics,

¹²³ Beyleveld and Brownsword. "Principle, Proceduralism, and Precaution in a Community of Rights." *Ratio Juris* 19, no. 2 (2006): 145.

¹²⁴ Beyleveld. "Human Cognitive Vulnerability and the Moral Status of the Human Embryo and Foetus." *The Contingent Nature of Life*, Ed. Marcus Duwell, Rehmann-Sutter, Christoph. & Mieth, Dietmar. (Springer, 2008): 85; *Law as a Moral Judgment*. (Sheffield: Sheffield Academic Press. 1994).

however, and while I believe his own method to be a dialectically necessary one, it is in Habermas's theory which we truly find emphasized the notion of epistemological uncertainty; indeed, it is epistemological uncertainty which lies at the core of Habermasian ethics and which does not so reside in Apel's theory, which is overtly foundationalist. We find this to be the case precisely because of Habermas's attempt at retaining a full rootedness in critical theory, with its dedication to the subjectivity and inter-subjectivity of truth. Thus, while Apel states quite clearly that what follows from a transcendental-pragmatic analysis of discourse is an ethics which acknowledges the primacy of "la reconnaissance mutuelle de tous les partenaires possibles comme ayant des droits égaux (au niveau de l'argumentation) et une égale co-responsabilité dans la découverte, l'identification et la solution de tous les problèmes moralement pertinents du monde de la vie par les méthodes procédurales du discours argumentatif," Habermas is explicit about excluding even any such talk of equal rights in discourse for fear of concretizing and rigidifying morality.¹²⁵ The main point I wish to make, however, is that I believe it is possible to reject the epistemological basis upon which Habermas attempts to justify such a non-concretization of morality yet to accept nearly wholesale the idea of epistemological uncertainty inherent in the resolution of interpersonal conflict. As Beylveled has demonstrated, it is possible and in fact necessary to accept that many moral dilemmas arise and persist because of such epistemological uncertainty. Through his derivation of the Principle of Precaution from the PGC, and through my present elucidation of the nature of the concept of epistemological uncertainty inherent in Habermas's theory, we begin to see the germs of a potentially powerful synthesis of Gewirthian and Habermasian theory. We see, namely, that the same sort of concept of epistemological uncertainty is emphasized in both theories, a concept of epistemological uncertainty which mandates that certain acts may not be undertaken with regards to other agents, and perhaps even other entities more generally, due to a sheer lack of information. Such a lack of information may result from an external contingency, as measured by the probability that some event or circumstance does not hold, or it may result from an internal contingency, as measured by the probability that, as an inherently flawed rational being, I am wrong or incorrect in my judgment. In either case, we shall now see, both conceptions of epistemological uncertainty defeat the aforementioned objection.

¹²⁵ Apel, *La Réponse de L'Ethique*. (Louvain: Editions de l'Institut Supérieur de Philosophie Louvain-la-Neuve, 2001): 77. "...mutual recognition of all possible partners as having equal rights (at the level of argumentation) and an equal co-responsibility in the discovery, identification, and the solution of all morally pertinent problems of the lifeworld via procedural methods of argumentative discourse."

3.4.4 *Epistemological Uncertainty and Alter-Contradiction: A Reconstruction of the Habermasean Argument*

A reply to the aforementioned objection can be offered in two parts, with the second part itself divided into two sub-parts. The first part of the reply consists in the fact that, due to the problem of epistemological uncertainty, I may not simply assume, even if I am a solipsist, that some entity is not an agent, when (1) all empirical evidence points to the contrary and when (2) it is just the case that I am a flawed rational being with the capacity for error in judgment. If I thus make the conclusion that an entity which is generally considered an agent is not an agent, that is I wrongfully reject the null hypothesis, then, as Beyleveld points out, I will have committed a gross violation of the PGC and will therefore have committed a highly immoral act. The same result holds on the foundationalist re-interpretation of Habermasean theory I am proposing: If I make the wrong conclusion, based upon either external or internal contingency, I will fail to abide by principles (D) and (U) and will therefore, on the Habermasean conception, have committed an injustice; on Apel's view, by contrast, this would entail a violation of the agent's rights. Beyleveld and others might still complain that the contradiction inherent in rejecting the conditions of the Ideal Speech Situation is not apparent. This thus brings us to the second part of the reply.

The second part of the reply likewise appeals to the idea of epistemological uncertainty. I cannot, on recognition of this fact, conclude that I am capable of arriving at a correct assessment of some fact and simultaneously conclude that the other is not capable of arriving at the same rational conclusions about this fact simply on the basis of the validity or soundness of the argument itself. This derives from the aforementioned fact of epistemological uncertainty inherent in the judgment of what entities are indeed agents. It so derives because, as stated elsewhere, if I act upon a norm whose fulfillment affects another agent (in terms of his negative or positive rights), and the agent does not accept that norm, I can act upon such a norm only if I conclude that that agent is not an agent, which is probably incorrect. This contradiction follows from the following set of assertions, which constitute the first of two sub-parts of the second part of the reply:

Premise 1: In prescribing any norm, I assume that any entity possessing logical judgment is capable of coming to the same conclusion, merely on the basis of the validity or soundness of the argument itself which yields the norm.

Premise 2: I implicitly assert that, since my norm is correct, its correctness can be discovered merely via the use of logical judgment. I thus assert that the only faculty necessary for the evaluation of the truth or validity of a norm is logical judgment.

Premise 3: If another Entity A disagrees with the norm, then it is probable that Entity A is an agent, on the basis of its having the capacity for disagreement or

agreement with the norm, which constitutes evidence of the possession of logical judgment.

– *Lemma*: I have strong evidence that Entity A is an agent.

Premise 4: Entity A possesses logical judgment if and only if Entity A is an agent.

Premise 5: In asserting a norm with which Entity A disagrees and proceeding to discover this disagreement, I implicitly assert or acknowledge, “Entity A does not accept the true or valid norm I have set forth.” Since Entity A possesses logical judgment, I implicitly assert “Entity A fails to use its logical judgment in the correct way.” Entity A makes the same implicit assertions with respect to me.

Premise 6: Since both Entity A and I are non-omniscient agents and are thus inherently fallible, it is possible that one of us is wrong.

Premise 7: The assertion that I am absolutely certain of the truth or soundness of my norm entails my implicit assertion that it is not possible that I am wrong with respect to this judgment. It follows from this that, if it is the case that I have arrived at these absolutely certain conclusions solely through the possession of logical judgment, then, if Entity A is an agent and thus possesses logical judgment, Entity A must arrive at the same absolutely certain conclusions.

Premise 8: Entity A does not arrive at the same absolutely certain conclusions.

Conclusion: Therefore, if I assert that I am absolutely certain of my conclusions and that I cannot be wrong about these conclusions, then I implicitly assert that Entity A is not an agent.

Thus, the contradiction here is not, as in Gewirth’s theory, self-contradiction or contradiction of one’s own agency; instead, it is what we may call alter-contradiction, or contradiction of another’s agency. From this conclusion, we see the germs of further conclusions, namely what I take to be the most important implication of Habermas’s theory: that, because I cannot be certain about the truth or validity of my norms, I may not impose my norms by force, for the imposition of my norms by force just is the most explicit assertion of infallibility, which I do not and cannot in principle possess. As early as 1793, William Godwin too seems to have hinted at such a link between the legitimacy of communicative action and the illegitimacy of the use of force. As he states:

Let us consider the effect that coercion produces upon the mind of him against whom it is employed. It cannot begin with convincing; it is no argument. It begins with producing the sensation of pain, and the sentiment of distaste. It begins with violently alienating the mind from the truth with which we wish it to be impressed. It includes in it a tacit confession of imbecility. If he who employs coercion against me could mold me to his purposes by argument, no doubt he would. He pretends to punish me because his argument is strong; but he really punishes me because his argument is weak.¹²⁶

If what I have concluded with regards to Habermas’s theory is correct, I believe that Godwin must be right and that, as he too concluded, freedom, and indeed a

¹²⁶ William Godwin, *Enquiry Concerning Political Justice and its Influence on Morals and Happiness*. (London: J. Watson. 1842): 161.

holistic and comprehensive sort of personal liberation, must be of very high value. Such a thesis, however, is far from being proved, and there is much in the way of theoretical elaboration that must be pursued before we may make such a conclusion concerning liberation or freedom more generally. It is at least interesting for now, however, that over 200 years later social cognitive theory has identified clear links between non-aggression and agency-affirming modeling conducive to effective self-regulation, especially as regards criminal punishment. As Bandura states:

Reasons combined with negative consequences generally produce stronger and more lasting self-restraints than negative consequences alone. When a transgressive act is punished, it is unclear whether that particular act or that general type of behavior is considered unacceptable. Discipline that is used as an occasion for explaining a rule of conduct is more effective in creating self-regulation of similar activities than if the specific act is simply punished. (Bandura, *Social Foundations*, 266.)

Let me, however, be clear. What I have been arguing does not imply that our norms are not, in principle, necessarily true *a priori*; it is not, in other words, our norms that are fallible (if those norms are indeed true) but instead we who are fallible in our judgment. Such a conclusion, while especially salient, is not the only salient one. As I shall show, the conclusion reached here concerning the premise of alter-contradiction at the heart of Habermasian theory yields more than merely a negative conclusion proscribing force against others; it yields a positive conclusion as well, one that also ends up giving us yet more tools for the resolution of seemingly intractable moral dilemmas.

3.4.5 *Applications to Moral Dilemmas: The Right to Discourse*

Beyleveld aptly points out that Gewirth's theory does not bear direct application to certain kinds of moral dilemmas, save through a general sort of proceduralism.¹²⁷ While I believe I have shown Gewirth's theory indeed to have much direct application to moral dilemmas, it remains true that it does not provide the tools for resolving certain kinds of moral conflicts and dilemmas because it simply does not provide a method, at least directly, of answering many of the very substantive questions that arise in such situations. As regards property theory, for instance, Gewirth's theory might tell us that we ought to establish a procedural method for handling such cases subject to the consent of the parties involved, but it tells us nothing about the actual normative content within such a formal framework. Gewirth's theory, therefore, might tell us what a legal system should look like in its general form, but it will tell us nothing about which parties within that system own which particular pieces of property and on what grounds they so own them. It

¹²⁷ Beyleveld and Brownsword, "Principle, Proceduralism, and Precaution."

will tell us whether we are to rescue or refrain from doing so, but it will not offer us any substantive reasons why we ought, for instance, to rescue someone even if it means a great cost or burden to ourselves. Lastly, it will even tell us, like Habermas's theory, that we ought to discourse with each other in order to resolve moral conflicts, but it will not tell us *exactly* how we are to interact with each other once in such a state of affairs. Habermas's theory, however, offers powerfully substantive answers to these questions, and while I will not undertake comprehensive analyses of such answers, I will elucidate the essential principles whereby Habermasian theory provides a resolution to these dilemmas. In lieu of comprehensive analyses here, I intend for such an elucidation merely to provide a basis for future inquiry into property theory, political theory, and other areas elsewhere. The questions I will explore with regards to Habermasian theory, as I have already hinted, will be three; these are three questions which I believe Habermasian theory, reconstructed in the way I have presented it, adequately answers. First, I will pose the general question, "On what grounds is it obligatory to enter into discourse?" I will answer that there exists what I shall call a right to discourse which is instantiated by particular kinds of circumstances and which, in addition, has definite determinate substance apart from being a merely formal procedural framework. Next, I will pose the question, "Who has a legitimate right to a particular piece of property when such property is not already clearly in any particular person's possession?" Habermas's theory, aided by an underlying virtue-theoretical framework, indeed offers a substantive answer to such a question. As it turns out, it is the right to discourse which is central to such a resolution. The final question I will pose is, "Is self-sacrifice permissible, and if so, are there any limits to self-sacrifice?" The answer I will offer will again rely upon the right to discourse, undergirded by a virtue-theoretical framework. These applications of Habermasian theory provide substantive answers to such questions, but, I will conclude, they cannot achieve this without the underlying structure provided by virtue theory. We shall see very explicitly why this is the case.

We have already seen, through my foundationalist reconstruction of Habermas's argument, how a right and a correlative duty to discourse might be established. As mentioned previously, Apel would probably be the most amenable to such an idea, as he himself has established that there exist certain transcendently presupposed rights rooted in the idea of rational communication. Beyleveld might also find such an idea as a right to discourse theoretically appealing, as, although perhaps not glaringly obvious, it is plausible that, upon further analysis, it could be shown that such a right is indeed derivable from Gewirth's theory and its commitment to proceduralism. Most recently, Rainer Forst has established a theory of justice founded upon what he calls "the right to justification," which is superficially similar to what I shall be proposing.¹²⁸ Unlike Forst, however, I will not be claiming that every person has a right to justification in the sense I believe he is using this phrase.

¹²⁸ Rainer Forst, *The Right to Justification: Elements of a Constructivist Theory of Justice*. (New York: Columbia University Press. 2012).

In other words, it is not the case, as he seems to imply, that individuals each have an equal right to *determine* what is just, in the epistemological sense of *constructing* truth. I will claim, rather, that it follows from a foundationalist interpretation of Habermasian theory that individuals each have the right and indeed the duty to *discover*, collectively, what is just. Of course, Forst is a student of critical theory, and so his constructivist interpretation is not an unexpected result.

I shall be highly specific with regards to what I mean by “the right to discourse” so as to assure the reader that I am not merely being a needlessly imaginative philosopher. I will begin the elucidation by posing once again the first aforementioned question: “On what grounds is it obligatory to enter into discourse?” It seems *prima facie* implausible that it could be obligatory always to be in a state of discourse or even to enter into discourse whenever one makes any sort of decision; it is implausible for the simple reason that it seems that such a situation could not even be carried out and would thus be subject to Kant’s rule that “ought” implies “can.” The simple answer to our question is that the right to discourse, as follows from my reconstruction of Habermasian theory, is instantiated whenever there is any sort of dispute, or even anything that could be called a disagreement or debate, in which each party to the disagreement has an equally legitimate right to make a decision with regards to the given state of affairs. The co-existence of equally legitimate decision rights with regards to some state of affairs thus instantiates the right of each individual to have equal input over how the state of affairs in question will be addressed. The clearest way we shall be able to see how such a right is applicable and how it thus resolves certain types of moral dilemmas is, as I have previously stated, through an analysis of property theory.

3.4.6 Applications to Property Theory

There are many diverging and conflicting theories of property and ownership, most of which are seemingly incompatible with each other and whose corresponding debates seem therefore to be equally intractable. While I cannot possibly do justice to property theory in this book, I will merely offer what I believe to be a very promising avenue for further inquiry. The idea of property, I believe, is a central concept around which revolve various pressing dilemmas, all of which I believe can, in principle, be resolved by appealing to Habermasian theory as I have interpreted it. One of the fundamental questions it seems we must ask is: “Who, before anyone owned property at all, had a right to a particular piece of property, and on what grounds did they have such a right?” Of course, on the right-libertarian or otherwise Lockean view, property rights lie in the hands of the individual who first finds the land, cultivates it, and creates something new from it. While I see much merit to this idea, there is also yet much unanswered as regards this position. For instance, what constitutes cultivation or development? What if more than one person at first finds this land? What is the logical connection between the newness of a creation or the concept of cultivation and rights to exclusive private-property ownership? What even constitutes *finding* the land? Need I have stayed on the land

for any appreciable amount of time before I own it, may I still own my property if I do not occupy it for most of the time, or must I stay there in perpetuity in order to retain my rights to it? Does my mere discovery of an entire uninhabited island or small land mass or even a body of water then grant me ownership rights to it? While I believe the concept of individual private property ultimately to be a legitimate one, these questions have not been adequately answered by right-libertarians. I believe the most important question concerning property theory, however, is not whether private property is a legitimate concept, for, even with these questions unanswered, the idea of private property might yet have legitimacy in a post-distributive and post-deliberative social situation. What I mean by this will shortly become clear.

These problems regarding the initial status of property ownership have impelled me toward the position that, in an idealized state-of-nature scenario in which all individuals have access to all resources and in which all individuals claim all resources, some indefinite amount of property is owned by each individual, but that it cannot be known, save via an initiation of discourse, which or how much property belongs to each individual. To be clear, by “to own” I mean “to have the right to hold in possession.” Such a right to property, it must also be furthermore emphasized, is not instantiated or *created* via discourse; rather, the right to a particular piece of property cannot be known or discovered without a process of discourse. Thus, the facts existed concerning what and to how much each individual had a legitimate claim, but these facts were simply epistemologically unavailable to the participants pre-discursively. This might seem a strange idea, but the theoretical element which clarifies this, as previously discussed with regards to the principle of proportionality, is that of *telos*. Once again, I must preface my statements concerning this concept with the qualification that it will be elucidated in great detail in later chapters. It is sufficient at this point, however, to demonstrate how, at least formally, the concept of *telos* grounds what I am claiming is a resolution to the moral dilemmas I discuss in this section. It is indeed the concept of *telos* which makes a Habermasian resolution to such dilemmas possible.

Let us revisit our scenario. We recall that my claim is that, because every individual initially claimed ownership over the entire extent of available property, and because every individual had an equally legitimate claim to all such property, every individual then had a right to decide upon how the land was used and even distributed privately. Initially, then, the land was subject to the rights of all, and all had rights to what all others did with the land. If it were then decided to give or to sell off pieces of land to private individuals, then those individuals owned the land, and all of the rights of private property that had formerly belonged to the group were then instilled, as regards this piece of land, in the now individual property owner. This, however, is a sort of idealized state-of-nature scenario familiar in social-contract theory, and so we must be careful to see exactly what I mean and to what, exactly, I am applying such rights. What I mean by “everyone owned all property” is that, given the moment at which agents emerge from a non-sapient state and can thus make voluntary decisions, and given that they all have access to the same finite amount of property (say, for instance, the entire earth), and given the

nature of positive rights to property derived from Gewirthian theory, it follows that no one person could claim a legitimate right to the entire mass of finite property, for it would deprive the legitimate positive rights to property possessed by the other individuals. It follows from this, then, that all individuals would each have an individual right to some minimal amount of this property and thus that, in order to ascertain what property in particular each person needs in order to fulfill this right, it would be obligatory that these individuals interact and communicate in at least some minimal way so as to come to some common decision regarding how property ought to be distributed. Even on a theory of merely negative rights, that is on a right-libertarian theory of property, it can be seen that, if each individual initially had just as much access to the entire earth as all of the others, then each would have an equally legitimate right to the entire earth, thus instantiating a right and a duty to discourse. The Habermasian model seems, then, to be at work here.

An important question, however, remains: "Does the assertion and effectuation of a right claim to some particular piece of property necessary for the fulfillment of rights to property necessitate, normatively, the principle of discourse?" Gewirth's theory, or at least Gewirth's interpretation of his own theory, tells us "no," while Habermas's theory clearly tells us "yes." For Habermas, in order for there even to exist a valid norm that affects or potentially affects anyone else, that norm must be accepted by those individuals affected or potentially affected, while, for Gewirth himself at least, the normative truth of the right claim to property, asserted, evaluated, and found to be true by the agent herself, is apparently sufficient for the justification of legitimate actions necessary to effectuate the norm. Is there some common ground between these two accounts with regards to this problem? I believe so. While it is true that it follows from Gewirth's theory that, in an atomistic sense, an individual has a right to pursue property necessary to himself or herself without any accountability to other individuals simply on the basis that the act in question undertaken is right and normatively justifiable and that, only if the action were unjustifiable would another individual have the right to intervene in some way, it is also true that, in the state-of-nature case previously described, two or more individuals might want to make use of or own the same property, be it a section of a stream, a portion of the forest, a place at the bottom of the ocean, or even the entire earth. Certainly, one can make the case, as Habermas does, that all other individuals affected or potentially affected by the decision must consent to the particular use of the property available intended by any one agent. Yet this does not contradict Gewirthian theory. Consider that, in this state-of-nature case, all have an equally valid right claim to literally the entire earth or otherwise finite stretch of property. In this situation, of course, it is clear how a particular individual's choice affects the others and thus interferes with the others' right claims; in such a case, we have stated, it is clear that the particular individual must abide by a group decision founded upon his own consent and that of the others. This, we stated, was a clear instantiation of the right and correlative duty to discourse. The question, however, remains: "But what if each individual chooses to take possession of property that cannot be seen to affect each other at all, save in extremely indirect ways?" I believe that we must conclude that, in such a case, there is no duty to discourse on the part of that individual, so long as these others do not have some disagreement with the individual with regards to how he is using his

property with respect to *them*. We must be clear, however, that the other individuals have no right to force the individual to discourse; instead, they have the duty of persuasion and incentive but not coercion. There is a more general sense, however, in which the individual has a duty to enter into discourse that is not directly related to the derivation of the right to discourse from Habermasian theory. This duty is separate from what I have called the duty to discourse in that it is a duty that is directly derivable not from Habermasian theory and interpersonal conflict but from virtue theory and individual *telos*. Again, I will cover such topics in greater depth in later chapters. Suffice it to say merely that, in general, individuals in the state-of-nature case have the duty to persuade the individual to interact, cooperate, and discourse for the purpose of the fulfillment of her *telos*, and the individual has the duty to so discourse for the purpose of fulfilling her *telos*, which necessarily involves the sharing of thoughts and experiences and which also necessitates a relation of non-coercion between individuals. Once again, however, both Gewirthian and Habermasian theory are clear that the others may not legitimately coerce or otherwise force the individual to discourse unless such a non-ideal speech situation were necessitated by the individual's violation of the others' rights (i.e., self-defense or other such similar circumstances). In other words, as the right-libertarian affirms, the individual does indeed have a right against being forced away from an isolated, atomistic existence free from nearly all social life and association with others, but the individual herself or himself actually does indeed have a duty to interact with at least some other agentive, rational beings in order to achieve her or his *telos*. This, of course, reflects left-libertarian insights, but it says nothing about property directly; strictly speaking, it is not obligatory, in order to achieve personal liberation, to share and administer common finite property with others; fulfilling and even friendly interaction does not necessitate such common ownership, unless of course the contingencies of the situation call for it (as in the case, possibly, of starvation or other similar situations). In the state-of-nature case, therefore, it is only in the case where the original property is disputed by two or more individuals that common ownership and management of the property via democracy is ethically obligatory; again, the right to discourse is instantiated. Once owned and managed by the claimants, the concept of a post-deliberative private property can be invoked if the claimants see it in their best interests and can thus legitimately sell off, give, or rent (as in a Georgist model) another individual a portion of the commonly owned property, having come to the decision via what can rightfully be called democracy. Let me now make one point clear: the very idea that two or more individuals may legitimately claim a right to some finite piece of property and thus instantiate a normative democratic ownership and administration of the property is derived from the fact that, while individuals have rights to a minimal amount of property necessary for individual *telos*, these rights are conditioned (and not conditional) in that individuals have the right to *a* portion of property but not *a particular* portion of property.¹²⁹ It follows from this that the individual has a right to *any* particular

¹²⁹ The conditioned-conditional distinction with regards to rights and duties is one established by Magnell.

unclaimed portion of property, which, in the state-of-nature case, is *all* possible portions of property necessary to satisfy the particular individual's right to property.

There are many critiques, however, mostly well-founded, of state-of-nature scenarios, so we must ask the practical question, "How do individuals come to legitimately acquire property in real life?" Let us assume a bit more realistic situation, if only a bit more. Let us assume that two parties exist on opposite sides of the earth (so as not to be in a situation where their actions even vaguely affect each other or where discourse is possible), and only these two parties exist. These parties also do not know of each other's existence, and these parties are also pre-industrial. Let us also assume they suddenly find themselves among unclaimed property and that party A is an individual and party B is a mass of individuals. Parties A and B, in this slightly more realistic case, do not collectively own all property (i.e., the entire earth) because they cannot claim exclusivity for, if each party does not know if other individuals exist which have already claimed parts of the earth, then each party does not know whether others have already claimed portions or indeed which portions have already been claimed. In this case, then, starting with party A, the only justifiable normative choice she or he can make is a Lockean one: choose a portion of land to use and hold (or if not land, certain objects or implements). Party B, however, appropriates property in a different way. There are at least two possible manners in which party B can legitimately appropriate property. The first is that party B sees itself as a community whose members decide to make decisions collectively and democratically. The second is that party B is comprised of individuals who coincidentally live together and simply happen upon certain property at the same time which they each individually claim as their own. In both manners of interaction, however, the normative result is the same: the individuals all collectively have rights to the property in the form of a right to discourse and thus have an effective right to democracy. Thus, as hinted before, the right to discourse with respect to property is thus truly a type of conditioned right to *some* property which fulfills the criteria of (1) not making the fulfillment of one's *telos* less probable and (2) positively rendering the fulfillment of one's *telos* more probable via the satisfaction of certain material conditions. Because both (1) and (2) are epistemologically unavailable to the participants save through discourse, the right to discourse is established. The reasons for the validity of such criteria, as well as the further reason why such criteria yield a right to discourse, will be explained further when we examine the next property scenario. For now, for the sake of argument, let us assume these criteria to be valid. Thus, no one in the group may thus make decisions about the property without the consent of all other individuals within the group. In this case, we observe an extension of the principles at work in the above-discussed, less realistic state-of-nature case which concerned the ownership of the entire earth, namely that, if two or more individuals find a piece of property at the same time, all of those individuals have a right to that property or, in other words, they all own the property. At this point, it is crucial to note that, although each individual finds the unclaimed property at the same time, it does not follow that they each have a right to *this* particular piece of property but only that they have a right to a piece of property that will satisfy their rights to property. The

key factor which instantiates each individual's right to this *particular* piece of property, however, is the very fact that, in having the right to *a* piece of property and a duty not to deprive others of their rights to *a* piece of property, it just ends up being the case that they may not drive each other from *this* particular piece of property by force. In other words, each individual has as much a right to *this* particular piece of property as every other individual among the parties who discover the property because of what right-libertarians call the Non-Aggression Principle, derivable both from Habermasian and Gewirthian theory.

3.4.7 *The Baker and the Starving Man*

The Non-Aggression Principle, however, in light of our derivation from Gewirth's theory of the Synthetic Double-Effect Principle, is in need of some reformulation. Indeed, although I believe the Synthetic Double-Effect Principle to be derivable directly from Habermas's theory as well, due to the theory's implicit reliance upon the statements agents necessarily assert with respect to other agents and thus on the concept of intention as I have presented it, I will not undertake a comprehensive proof of this here. Instead, I will merely submit to the reader what follows from the Synthetic Double-Effect Principle with regards to the Non-Aggression Principle, given the soundness of Habermas's theory. As we saw with regards to Gewirth's theory, the Synthetic Double-Effect Principle is entailed by the PGC. As a result, we concluded that the following principle holds:

The Synthetic Double-Effect Principle: *Never aim or attempt to maximally violate agents' rights to freedom, and always aim at or attempt to not-maximally-violate agents' rights to freedom, even if it is in order to maximally uphold agents' rights to freedom.*

Thus, with respect to the phenomenon of aggression, we concluded, it follows that one must do only what is necessary to uphold one's rights to freedom without aiming or attempting to violate any other agents' rights to freedom and while simultaneously aiming or attempting to uphold all other agents' rights to freedom. We saw this principle in action most explicitly in our analysis of self-defense. This same principle, we must see, derives directly from Habermas's theory, and indeed, in order to act with respect to any other agents in the Ideal Speech Situation, it is necessary to follow this principle. Indeed, we begin to see, at this point, that the Principle of Synthetic Double Effect, on a freedom-centric interpretation, just is the Non-Aggression Principle. Let us be clear, however, what this principle implies. On the standard right-libertarian interpretation, the Non-Aggression Principle prescribes for us the directive not to employ force, save in cases of self-defense. Of course, this principle is based upon the fatally flawed premise that there exist only negative rights. Thus, what would follow from the principle, "One must do only what is necessary to uphold one's rights to freedom without aiming or attempting to violate any other agents' rights to freedom and while simultaneously aiming or

attempting to uphold all other agents' rights to freedom" is precisely this result. If, however, we are to conclude that positive rights exist, in other words that "rights to freedom" includes negative rights to freedom as well as positive rights to freedom, then we are compelled to accept the logical result, namely:

The Non-Aggression Principle: *One must do only what is necessary to uphold both one's negative and positive rights to freedom without aiming or attempting to violate any other agents' negative or positive rights to freedom and while simultaneously aiming or attempting to uphold all other agents' positive and negative rights to freedom.*

Thus, while on the standard right-libertarian interpretation, we would conclude that the starving man about to die from hunger would unjustifiably violate the rights of the well-to-do baker in stealing a loaf of bread from him, on the present interpretation of the principle, we would have to admit that, while the starving man engaged in an act that had a bad effect, namely the violation of the baker's rights to his property, the starving man was nevertheless justified in such an act because (1) all other choices entailed a more gross violation of rights to freedom and (2) the starving man attempted to do everything in his power *not* to violate the rights of the baker and positively to uphold the rights of the baker (i.e., by discoursing with the baker, perhaps in the form of begging, or by attempting to do everything possible, within permissibility, to obtain money to buy the bread instead). Thus, although Habermas's implicit non-aggression principle seems superficially to confirm, for instance, Hoppe's right-libertarian view, it is clear, on closer examination, that although the starving man does assert infallibility against the baker and thus contradicts the baker's agency in so acting against the baker, this is not his *intention*, as I have characterized it, for, since the dilemma involves two parties, claims of infallibility flow in both directions. This is evidenced by the fact that the starving man did everything in his power, as well as within the limits of permissibility, *not* to presume such infallibility as regards the baker. As the Synthetic Double-Effect Principle demonstrates, the violation of the baker's right was a mere by-product of the starving man's act and in fact was contrary to his aim to uphold his own right. Thus, while each party in the Ideal Speech Situation must not assume her own infallibility and thus impose her norms on the other, she must, as a matter of practical necessity, at the very least accept the conditions of the speech situation of which she is a part. She must, in other words, as all agents implicitly do, accept as valid at least the very basic norms and rights entailed by the Ideal Speech Situation, that is, if she is to act voluntarily at all. Thus, she may not presume *nothing* lest the phenomenon of her voluntary action cease altogether. Thus, we must be clear that what follows from the Habermasian conclusion I previously set forth is not that we must not presume any moral truth but, rather, that we must acknowledge the fact that we might be wrong. The basic norms derived from the Ideal Speech Situation (and we might as well say from Gewirth's theory also), however, are the best she can judge, and if she finds herself in a situation in which her choices are so limited that the exercise of her abidance by these minimal norms is placed in jeopardy, she has normative latitude only to make the choice that entails

the least unjustifiable normative assertion with respect to those norms. In other words, while she must acknowledge that she might be wrong about her norms, she must, contrary to denying the agency of the other, likewise acknowledge that the other agent is, *qua* agent, fallible and, thus, might also be wrong.

Thus, we must see, it follows from this analysis that we must recognize not merely one but two kinds of aggression: the right-libertarian conception and the left-libertarian conception. While it is certainly the case that the failure on the part of an agent to respect a negative right is wrong and is what we are calling aggression, it is also wrong and also an aggression, if we accept the existence of negative rights as they are derived from both Gewirthian and Habermasian theory, for an agent to fail to respect a positive right. A choice, we must recognize, is indeed an action. Thus, if we acknowledge that the choice of one will affect the freedom of another, there is no reason why we should distinguish between negative and positive forms of freedom; indeed, Arrow's Impossibility Theorem itself seems to undermine this distinction. This, of course, is a classic left-libertarian insight and the basis upon which left-libertarians and left-authoritarians alike have critiqued liberalism, namely that liberals focus solely on negative freedom and thus neglect the power structures that exist which limit and restrict positive freedom. We may analyze the individual choices that people make in a liberal society, they assert, but in doing only this we will overlook the relevance of those choices made by those in power. It is the choices made by those in power, we must realize, that are of utmost normative relevance, for it is precisely those in power who possess the greatest capacity to make choices which limit the positive freedom of others.¹³⁰ Thus, if she presumes that she might be wrong and that the other might likewise be wrong, then, upon realizing this, the only thing she may do is aim or attempt to uphold the norms she holds while making every possible effort, within permissibility, to avoid imposing her norms on the other party. The baker must do likewise. We must see, however, that in failing to take the interests of the starving man into consideration, he indeed does impose his norms on the starving man. In failing to attempt to abide by the conditions of the Ideal Speech Situation, he fails to be accountable to the starving man and thus to persuade him or otherwise gain consent for acting toward the starving man in this way. Thus, in not giving the starving man his loaf of bread, and in not having the starving man come to the same conclusion regarding this matter, the baker acts as if his norm were infallible. Each party, we see, must make some presumption of infallibility to act in the aggressive manner that he does, but the truly relevant question here regards which act is more justifiable, or, in other words, least unjustifiable. If, in the final estimation, the starving man finds himself faced with two unjustifiable choices, namely choosing to steal or choosing to starve,

¹³⁰ While the solution that left-authoritarians, and most notably Marxists, seek to such a power asymmetry is the imposition of their own power asymmetry via the establishment of new structures of elite power and domination, left-libertarians seek solutions in the equalization of discursive power via collective resistance of such structures of elite power and domination. Interestingly, the analysis thus far clearly precludes the legitimacy of the Marxist or otherwise left-authoritarian approach.

and both are proscribed by the Ideal Speech Situation, then, to his knowledge, the most accurate conclusion he can make, given the circumstances, is that he ought not to starve.¹³¹ This sort of situation is, for sure, highly unjust, but we must see that, in such a situation, it is literally the best judgment the starving man can make.

3.4.8 The Right to Discourse, Self-Sacrifice, and the Antique Clock

At this point, I have spoken much about what I have called the right to discourse, but this concept may be critiqued on the following grounds. One might argue that Habermas, in never having conceived of discourse in rights-based terms, primarily conceived of discourse as standing in as a specifically epistemological resolving force of moral conflict and so would not have accepted what might be seen as the subsidiary role I have given it. In this light, it might also be argued that my refiguring of Habermas's idea of discourse is circular, for if Gewirth's monological process *ipso eo* resolves moral questions between individuals, then why do we need a right to discourse? In other words, individuals may have the right and duty to discourse in particular situations in order to resolve moral conflicts between each other, but once in the Ideal Speech Situation, what arguments should they actually make with regards to each other? Should they use Gewirth's theory, and if so, why was discourse then necessary in order to epistemologically determine the rightness of a given course of action anyway? My use of Habermas's theory through the so-called right and duty to discourse, then, seems at best vacuous and at worst circular. This, however, is not the case. Let us examine the cases I have presented in which I have claimed the right and duty to discourse applies. With regards to property theory, I have claimed, for instance, that the right to discourse applies in democratic groups who collectively own and administer property, as well as to two or more parties who simultaneously discover and claim a piece of property. One may claim of these two cases that the so-called right and duty to discourse is superfluous due to the monological nature of Gewirth's theory; in other words, all possible legitimate claims can simply be derived from Gewirth's theory without having to resort to Habermasian concepts. This claim, however, is deeply problematic and demonstrates just the sort of limitations which exist in Gewirthian theory. While it is true that it can be determined formalistically and *a priori* what basic rights a particular individual has (for instance, to food, water, and shelter), Gewirth's theory does not and cannot in principle reveal *a priori* to whom a particular antique clock or beat-up suitcase belongs, that is, to whom particular items belong which are either unclaimed or mutually claimed. As regards the issue of self-sacrifice, the same sort of complaint is applicable, namely that Gewirth's

¹³¹ See here that I emphasize that he makes the decision that he ought not to starve, rather than that he ought to steal. There is a substantive difference here as regards his intention.

theory does not take into account the subtleties of interpersonal interaction which turn out to be highly relevant to the concept of *telos*-fulfillment and, thus, highly morally relevant. My basic argument as regards these cases will be that Gewirth's theory does not actually resolve these dilemmas and that it is Habermas's theory and virtue theory which fill in the missing pieces. These dilemmas can be resolved only, I will argue, discursively.

I will work backwards so as to more clearly reveal the principle I am trying to elucidate. The scenario which best elucidates Gewirth's theory as it applies to the issue of self-sacrifice is one, first cited in *Reason and Morality*, in which a person is drowning who, for whatever reason, is capable of being saved only by me. Gewirth claims that, so long as the harm done to me in saving person A is not *equal* to that done to person A in my not saving him, I have a duty to save him. This conclusion, however, has weighty consequences. Thus, whether I am permanently disabled, financially ruined, or psychologically traumatized, Gewirth implies that, so long as I do not lose my life, which is what would happen to person A if I do not intervene, I have this duty. This, I believe, is far too simplistic. Gewirth makes this argument on the basis of what rights in general are of greater necessity for freedom and well-being. This concept of "degrees of necessity," as I have stated before, is a problematic one, but we shall, for the sake of the argument's elucidation, retain this concept. In later chapters, I will replace this concept with one which I believe to be more accurate. What is truly relevant to such a scenario, I believe, is a consideration of which action will result in the promotion of the *summum bonum* and, in addition, which action, if undertaken, will result in the impossibility of the fulfillment of the *summum bonum*. When we frame the scenario in this way, we are faced with an obvious question: "Which possible outcomes in this situation are most likely to lead to the fulfillment of the *summum bonum*?" It is here that the answer is less obvious than one might expect. Again, using a virtue-theoretical framework which I will later comprehensively analyze, it is quite obvious, in the case of the drowning person A, that the *summum bonum* will be made impossible for him by his death. But it is at this point, if we take Gewirth's account of this situation to be sound, that we seem to totally neglect the would-be rescuer. We do not, for instance, ask whether the attainment of the *summum bonum* would be made impossible for him *by* the act of saving person A. Gewirth, as well as most ethical theorists, seem to overlook this question because most of them do not conceive of *telos* as individual (and many do not conceive of a *telos* at all). In later analysis, we shall see that, although there is a general *telos* which encompasses all other possible *teloi*, there are indeed still individual *teloi* that exist as subclasses of the larger *telos*. We must not forget about these *teloi*. Thus, when we ask the question, "Which outcome is most likely to lead to the fulfillment of the *summum bonum*?" we must consider that it is possible that, without dying or even coming close to dying, the particular incarnation of the *summum bonum* that is the would-be rescuer's *telos* could indeed be made impossible by the act of rescuing person A. If indeed it is the case that the act of rescuing person A would make the would-be rescuer's *telos* impossible, then it cannot be virtuous or good, in this situation, to save person A's life because it would be inconsistent with the maxim that one must act to promote the *summum*

bonum while not intending and intending not to make the attainment of the *summum bonum* impossible. What does this have to do with the principle of discourse or even with the right and duty to discourse? We shall soon see.

Let us analyze the first situation I mentioned, that of the antique clock or beat-up suitcase. An antique clock seems personally more appealing, so let us have that as our example object. Certainly, a particular antique clock is not necessary for my survival and thus is not necessary for the fulfillment of my basic rights to food, water, and shelter; it is superfluous in this regard. Leaving aside the more fundamental rights, then, what can we say about who legitimately owns the antique clock? In order to determine ownership of the clock, we must see, Gewirth's theory does not help us, for, at least directly, it addresses only these more fundamental rights with respect to property ownership. Habermasian theory, paired with my own account of virtue theory, however, provides an answer. Let us imagine that two different people, person A and person B, claim to own the clock. Who owns this clock? Again, the ultimate and ultimately legitimate appeal is to individual *telos* and the *summum bonum*. The first question we must ask, similar to before, is "Does the absence of the holding of the clock in person A or person B's possession make person A or person B's *telos* impossible?" If one of their *teloi* is made impossible by the lack of ownership of the clock, then it is best for the person in question to hold the clock in personal possession. It follows from this that the person in question has a right to that clock and, more specifically, a property right to it. This seems a highly unlikely scenario, but it is at least conceivable. The second question we must ask is: "The fulfillment of whose *telos* is made more probable by the personal possession of the clock?" Again, whoever's *telos* is made more probable by the possession of the clock has a right to it, and more specifically a property right. For instance, the clock could be an heirloom from a former Holocaust survivor whose presence in the person's life inspires him or her to do and achieve certain acts or states of mind that might be more difficult to achieve otherwise. But force, we must remember, is a great evil. This implies that, although these property rights may exist, their fulfillment must not be gained by force unless the risk of failing to fulfill some other more necessary right justifies the use of force. Because force is generally illegitimate and evil, therefore, if another party wishes to acquire the clock, it must be acquired via communicative action within the Ideal Speech Situation (or via some deliberative procedure that both participants agree to, which in turn must also be consistent with the conditions of the Ideal Speech Situation). The instantiation of the Ideal Speech Situation, and thus the right to discourse, is necessary precisely because, once again, there is an epistemological problem seated at the heart of our ethical problem. One may retort that a monological method, whether dialectical and Gewirthian or assertoric and aretaic, can adequately address this situation, at least in principle, because it is theoretically possible, as I have supposedly just demonstrated, that an agent can individually determine, within the confines of its own mind, whether the possession of the clock will make the fulfillment of its own *telos* more probable. This, however, is patently not the case. The epistemological problem cannot be resolved via a Gewirthian monological method because, in such a case, it is clearly the case that, until one engages in some sort of discourse with the

other claimant or potential claimant, one cannot know, until one hears the other's reasoning, whether possession of the clock will make the fulfillment of their own or the other's *telos* more probable. Simply upon analyzing the very concept of probability, one cannot know what is more probable without having at least two cases to compare. In this case, the only way that comparison of such probabilities is possible is through discourse. Nothing can in principle substitute for discourse either, for only through an elucidation of the other's experiences and self-analysis can one even begin to determine the probability that the possession of a particular object will fulfill another's *telos*. We shall see the truth of this conclusion more clearly in later chapters, where the essence of the excellence of experiential rationality is thoroughly elucidated. Thus it is in this way that Habermas's theory, and specifically Habermas's theory, offers potential resolution to the most nuanced moral dilemmas involving the most nuanced of interpersonal interactions. We must see, however, that such potential resolutions are only possible through a collaboration with post-dialectical, assertorically necessary virtue theory. As above stated, this point will become clearer as we progress.

As for the question concerning Gewirth's drowning victim, post-dialectical, assertorically necessary virtue theory also offers a more extensive and nuanced resolution. On Gewirth's view, the potential rescuer has a duty to rescue the victim because saving the drowning victim will not result in the *de facto* violation of the rescuer's right to life, which is precisely the right which will be violated with respect to the victim if the victim drowns. Once again, however, we must (again as will be shown extensively further on) consider each individual's *telos*. While it is indeed true that the victim's necessary preconditions of agency will be extinguished if she or he drowns, as we shall see, the transcendental-pragmatic recognition of the necessary preconditions of agency are not the *sole* criterion which determines the rightness or *goodness* of an act, the semantic substance of which I have been claiming Gewirth's analysis lacks. What fundamentally determines the goodness of an act is action in accordance with the *summum bonum*, manifested in individuals as individual *teloi* corresponding to respective individual essences but which ultimately derive via a fundamental set of relations from the *summum bonum*. Thus, when the potential rescuer sets about making a decision regarding whether he will save the drowning victim, he must ask himself, in the first instance, whether he has, via the PGC, a *prima facie* duty to save the drowning victim. If such a *prima facie* duty exists, which it does, the potential rescuer must then proceed to consider whether his act of saving the victim will incur any harm to himself in the sense of making his fulfillment of his *telos* impossible. If the act of saving the victim does indeed make such a personal fulfillment of *telos* impossible, perhaps via some irreversible injury (physical or psychological) or some other severe harm, then, in this light, we can see that the maximal violation of one's rights to the generic conditions of agency, while not incurred by either agent in this scenario if the potential rescuer does rescue the victim, is not exhaustive of the criteria relevant to resolving this dilemma. Instead, as we have seen, we must appeal to a more fundamental question of individual *telos*, and in a manner very much in the spirit of Gewirth's original account. Whereas Gewirth cites the generic conditions of

agency as supremely relevant, I cite the *summum bonum*, and thus the individual *telos*, as supremely relevant in that, although performing a good act might not result in the extinction of the agent's necessary preconditions of agency, it might well be the case that such an act results in the extinction of the possibility of that agent's fulfillment of his *telos*. At this point, two questions in particular might arise to tantalize the reader. The first concerns why *individual telos*, as opposed to perhaps communal *telos* or even something like national or historical *telos*, serves as the resolving factor in such dilemmas. The second concerns the epistemological determination of individual *telos* or, stated differently, how one would go about discovering one's *telos* in the first place, let alone determining whether or not one's *telos* is made impossible by a particular act. These, of course, are contentual questions, and we shall answer them in later chapters. My main point here is to demonstrate that, if my theory is sound, then such a method of resolution of moral dilemmas follows from Habermasian theory. In any case, we must see once again that, at least in similar scenarios where the impossibility of *teloi* between parties is more ambiguous (it is not at all ambiguous between the parties in this scenario, since the victim will die) the right, as well as the correlative duty, to discourse arises due to epistemological uncertainties concerning two or more *teloi* that only discourse between the parties can render more certain and less ambiguous. This right to discourse, foreign to Gewirth's and even to Habermas's theory (although perhaps not to Apel's theory), is one which, in appealing to broader questions concerning value via the concept of *telos*, and not simply to duty and justice, fills certain gaps of theoretical applicability in Gewirth's and Habermas's respective theories.

3.4.9 *Historical Arguments for a Reinterpretation of Habermasian Ethics*

In *Habermas: An Intellectual Biography*, historian Matthew Specter explores Habermas's role in the politics and law of West Germany and attempts to explain much of the fragmentary nature of his normative views.¹³² Some of this, he argues, can be explained by Habermas's desire to overcome existing intellectual discourses in West Germany. "Habermas's early writings," he explains, "are a daring—even jarring—hybrid of existing discourses. Habermas's interdisciplinary reach seems to have been driven by the peculiarities of the West German intellectual field, which was structured by an empirical and positivist political science on one hand, and a conservative statist constitutional theory on the other."¹³³ Habermas became distressed by what he thought was an increasingly technocratic and scientific approach to politics. As historian Timothy McCarthy explains of Habermas's views: "Both decisionistic and technocratic models of political practice reflect the

¹³² Specter, *Habermas*.

¹³³ *Ibid.*, 60.

transformation of practical into technical questions and their consequent withdrawal from public discussion. In the former the citizenry serves only to legitimate ruling groups by acclamation through periodic plebiscites. Political decisions themselves remain beyond the authority of rational discussion in the public sphere; decision-making power can be legitimized but not rationalized. In the technocratic models, on the other hand, the exercise of power can be rationalized but once again only at the expense of democracy.”¹³⁴ For Habermas, this was the “reduction of political power to rational administration” and not, as he had hoped, a more broadly participatory regime in which democracy flourished in all areas of social life.¹³⁵

The legal system, as Specter explains, presented one of the greatest obstacles to democracy for Habermas. While other contemporary commentators described the Nazi regime as one of legal positivism, Habermas recognized that this regime was anything but positivist. While Habermas early on abhorred legal positivism, he described the Nazi regime as adherent to a strong form of natural-law theory. As such, it placed a heavy emphasis on the values of Nazism and the Nazi legal apparatus. This dogmatic appeal to value systems as a foundation for legitimate politics and law was exactly what he saw in the new Federal Republic, only this time the dogma was parliamentary democracy and not fascism. “Habermas’s insistence on a strict distinction between norms and values,” Specter explains, “responds to the hegemonic discourse of common ‘values’ introduced by the post-World War II Federal Constitutional Court as a means of integrating West German society. Its so-called value jurisprudence was dominant from the early 1950s through the early 1970s and was shaped decisively by the Weimar-era writings of jurist Rudolf Smend. . .” He continues: “By conceiving of fundamental rights as ‘basic values’ that expressed a social consensus on the notion of ‘the good,’ Smend’s value jurisprudence elevated the highest West German court to a position of political power that was, in Habermas’s view, beyond the reach of any democratically legitimated legislator.”¹³⁶ As such, Habermas developed a deep mistrust of the West German legal system, as well as the notion of a judicial branch and judicial review more generally. For Habermas, then, true democracy would have to shed itself of any claims to ultimate foundations in values. It would, in short, have to ground itself in political theory guided by morality, that is, by the notion of justice. For Habermas, Specter argues, “only theoretical reconstruction could rescue the insights of republicanism and liberalism from their too concrete embodiments.”¹³⁷ The democratic constitution should thus always be fluid, dynamic, and changing, and democracy should expand ever more into new spheres of social interaction. The *telos* of this process would be utopia. One prominent explanation, then, for the critical-foundationalist tensions we see in Habermas’s ethical theory might also be due to his experiences with the plebiscitary, technocratic system of

¹³⁴ McCarthy, *The Critical Theory of Jürgen Habermas*, 11.

¹³⁵ *Ibid.*

¹³⁶ *Ibid.*, 196.

¹³⁷ *Ibid.*, 175.

West Germany. In order to fend off what he perceived as threats of force and distortion on the part of the State, it was necessary to continue in the anti-foundationalist tradition of critical theory while simultaneously finding a new way to stand his theoretical ground by incorporating strains of foundationalist thought into his normative framework. By doing so, however, he would dramatically challenge some of the fundamental tenets of critical theory itself.

This historical context may ultimately help us to reinterpret Habermas's theory either along more strongly critical-theoretical or foundationalist lines. What is certain, however, is that Habermas cannot place a strong emphasis, as he currently does, upon both. This, however, entails two separate dangers for Habermas. If he wishes to keep communicative ethics within the realm of critical theory, it is likely that he will have to revert to previous modes of normative interpretation from which he purposively attempted to retreat in *The Philosophical Discourse of Modernity* and other works.¹³⁸ He might have to accept, for instance, the inherent oppressiveness of reason or its inherently destructive and self-destructive nature. He might additionally have to accept its supposedly technocratic and scientistic tendencies. In this case, then, his project of communicative ethics might exceed the degree of fluidity and indeterminacy that Habermas set out to establish. Communicative ethics, in fact, might completely disappear. If he wishes to transition to an openly analytic or otherwise foundationist model, however, he faces the risk of establishing an ethics with far too much concreteness that might itself result in oppression in the name of values. It seems, then, that Habermas is stuck in a similar aporia concerning reason that plagued the early critical theorists. Although this tension definitely exists within Habermas's normative theory, Habermas might yet have a way out. While most of the philosophical questions with regards to earlier critical theory have focused on the irresolvable aporia that Horkheimer and Adorno established concerning the impossibility of liberation due the self-destructive and oppressive nature of reason, it seems that an altogether different question has been neglected, namely "What are oppression and liberation but normative concepts?" Questions concerning oppression and liberation are both literally and phenomenologically meaningless without the implicit acceptance that reason frames and indeed normatively supports the entirety of such discourse. To say that reason is inherently oppressive, then, is not only performatively but also deductively contradictory. As Habermas has argued throughout his works, critique requires more than mere negation; it requires synthesis, a positive use of *Vernunft*. Out of Habermas's theoretical tension, then, comes a fundamental truth: the continuous effort to escape from reason is not a liberatory process; it is instead part of the problem. While the earlier critical theorists saw in rationality and critique a destructive force deleterious to liberation, a radical revision to this idea is long overdue. Habermas's normative thought, it must be seen, contains the seeds of such a revision. Quite the opposite, critique, even if of rationality and liberation themselves, is inherently liberatory and has the effect of preserving states of liberation where they already

¹³⁸ Habermas, *The Philosophical Discourse of Modernity*.

exist by continually placing and re-placing paradigms into the sphere of discussion. Rationality, as I have claimed from the beginning, leads the path to liberation.

3.4.10 *Concluding Remarks: The Single-Person Problem*

Our collective discussions throughout this chapter have thus far enabled us to make some firm conclusions. We have thus far demonstrated (1) that non-descriptivism cannot be sound and that some form of naturalism must be sound; (2) that, through the exposition of a negative model of ethical justification, there are certain attempts at justification that are unsound; (3) that, through the exposition of a sound positive account of ethical justification, certain determinate directives for action are derivable from purely naturalistic premises; and (4) such a sound positive account has far-reaching application to real-life, as well as unlikely, scenarios. There are, however, some issues that remain. We have defined neither “good” nor “ought” and its related deontic terms, and neither Gewirth’s nor Habermas’s theory provides us with such information; they provide us only with the analytic truth that agents must make certain deontic assertions and assertions of goodness. We must judge our freedom to be good, we must act for some purpose, we must judge that we ought to do certain actions, and we must make value judgments. But what is “must,” and what is goodness?

Another problem in Gewirth and Habermas’s theories is what Slote calls the problem of “self-other asymmetry.” He does not explicitly address Gewirth or Habermas but, rather, all deontological, and eventually utilitarian and consequentialist, theory in general. This problem, he explains, is one which concerns the degradation and devaluation of the agent:

I do want to claim that these forms of morality downgrade the (actual) importance of moral agents and their individual concerns, projects, even desires. But the argument for this claim will not be that, like consequentialism, the ordinary and/or the Kantian standard of right and wrong make an insufficient concession to the moral agent’s welfare. On the contrary, it will involve, rather, the claim that common-sense and Kantian morality *do* make concessions to the well-being and happiness of agents, *but make them only as concessions*.¹³⁹

Some theorists might quickly brush off Slote’s comments, particularly because they might perceive that deontological, consequentialist, and “common-sense” principles and rules provide all of the duties, concern for one’s own happiness, or concern for one’s own desires that are necessary for a comprehensive ethical theory. “Who cares,” they might ask, “whether these are ‘concessions’ or not? They still provide a basis for self-concern.” Slote’s comments, however, are not so easy to brush off. In order to demonstrate Slote’s point more clearly, I offer deontologists, consequentialists, and others what I shall call The Single-Person Problem. All of these theories, including consequentialism, derive their principles

¹³⁹ Michael Slote, *From Morality to Virtue*, (Oxford: Oxford University Press. 1992): 13.

and rules from social interaction, that is, from a set of relationships agents logically or otherwise have to each other. But what if you are the “single person?” Some might scoff at the following demonstration on grounds of its high improbability. Again, improbability should not matter a bit. If we are to truly test and vet our theories, they must be subject to *logically possible* thought experiments. It is logical possibility and not improbability that truly matters.

What, exactly, do I mean by the “single person?” Imagine that a series of catastrophic nuclear blasts killed everyone on earth, that is, everyone but you. Imagine, also, that all animal life was also destroyed; all that is left is you and some vegetation. It is at this point that we must seriously question whether any deontological, consequentialist, or any such principle- or rule-based theories retain their practical application. Deontology, of course, revolves around the concept of duty, duty which is instantiated by the existence of others, whether they be agents or non-agents. Where, within any deontological framework, can we find a principle, usually based upon the universalization of some maxim, that does not ground its legitimacy in the moral relevance or existence of other beings alone? Any deontological contractarian theories, most notably those of Rawls and Nozick, break down, for there no longer exists the possibility of coming together under a “veil of ignorance” and forming a conception of justice that directs the actions of all persons. Similarly, all variants of the contractarian approach, including Scanlon's inverted reformulation and even Parfit's reformulation of Scanlon's reformulation, break down. Habermas's attempt to remedy contractarianism by eschewing veils of ignorance or tacit agreement in favor of direct discourse also fails, for with whom can we enter into discourse? Even if we changed the situation to include the continued existence of non-agent animals, these frameworks would still fail since all of their principles and rules are derived from social interaction and relationship alone.

While Slote acknowledges that consequentialism avoids many of the same problems as deontology in this regard, it eventually falls prey to problems of its own. Consequentialism, he states, at least accounts for the self in some broad way such that one's own happiness is added into the calculation of overall happiness. As he notes, however, consequentialism lacks a firm self-regarding basis, particularly because consequentialism obliges one to sacrifice one's own happiness, or other value, if it is necessary for the best overall consequences or, specifically in utilitarianism, the greatest overall amount of happiness. Consequentialists are, then, still bound by relevantly similar standards of self-other asymmetry.

The main issue, as the reader might be beginning to realize, is that self-other asymmetry results in these theories' lack of application to all scenarios, namely the single-person scenario. The precise problem is that, when it is only the single person that exists, we can no longer account, via these theoretical frameworks, for ethical judgment. Ethics, it seems, is no longer applicable. No standards of behavior can be established, no reasons for action can be given, and no account of ethical judgment can have any grounds on which to stand. The single person stands alone, wayward, wandering, and without direction. Even Gewirthian rationalism accounts only for what an agent *must* assert in order to be an agent, and these

assertions have no binding force whatsoever unless there are other agents. Even the duties to oneself which are derived from Gewirth's theory are derived from judgments to which we must assent concerning all agents, namely "I must uphold the rights of all agents," which entails "I must, or have a duty to, uphold my rights." Ethics, it seems, mysteriously disappears. What, then, of ethical theory? It is through the Single-Person Problem that we see the grave consequences of relying on deontological or consequentialist frameworks alone. If any of these frameworks are to be sound, there must be a more fundamental, sound framework underlying them. As we have previously seen, the negative method of justification demonstrates that there remain, despite the existence of the single person, certain attempted justifications one cannot use and certain reasons one cannot give for one's choices. Even in the single person's world, then, there are such things as invalid and unsound reasons for action. If there exist such reasons, and naturalism is a sound framework, then it is not inconceivable that we can still find natural criteria upon which we can ground good or justifiable, even if perhaps not "moral," action. In his article, "Rights and Virtues," Gewirth tells us that, if virtues are to have legitimate standing in ethics, they must have a firm grounding in agent rights.¹⁴⁰ I hope that, through my exposition of the Single-Person Problem, I have made it obvious that we have rather solid grounds for making a different conclusion. As we shall see, in order for such concepts as rights and duties to have legitimate standing in ethics, they must have a firm grounding in virtue theory. Proving this, however, necessitates a deeper analysis of deontic concepts such as "ought," "must," rights, and duties. As we shall see in Chap. 5, these concepts are inextricably tied to and dependent on the concepts of goodness and virtue. For now, let us take a deeper look at the deontic concepts themselves.

¹⁴⁰ Alan Gewirth, "Rights and Virtues," *The Review of Metaphysics* 38, no. 4, (1985): 739–762.

Chapter 4

The Dialectical Structure of Value Judgments

In this chapter, I utilize the dialectical structure of value judgment provided by Gewirth and Habermas's theories in order to show that deontic concepts and their correlative judgments are logically and semantically intertwined with aretaic concepts. Importantly, however, deontic concepts are necessarily dependent upon aretaic concepts for their content and meaningfulness.

4.1 The Dialectical Structure of “Ought” and “Must”

In this section, I analyze the deontic concepts “ought” and “must.” I demonstrate that they have very direct logical connections to value judgments containing the terms “good” and “value.” “Ought” and “must,” I furthermore argue, are principally distinguished by their differing modes of conditionality and subjunctivity.

4.1.1 *The Harean Account of Imperative, Indicative, and Conditional Statements*

“I ought to have freedom,” “I should have freedom,” “I must have freedom”: What do these assertions actually mean? While much work has been done on the analysis of “ought”-judgments, most of this work has been done only within the context of the Is-Ought Problem, and its relationship to goodness is almost never analyzed. This, unfortunately, is probably due to the vast amount of non-descriptivist thinking in ethics. Analyses of the Is-Ought Problem abound, including, of course, those of Hume,¹

¹ David Hume, *A Treatise on Human Nature*, (London: Longmans, Green, and Co. 1882).

G. E. M. Anscombe, Philippa Foot, Geoffrey Hunter, Hare, W. D. Hudson,² Gerhard Schurz,³ John Searle,⁴ and especially Gewirth,⁵ among numerous others. So much work has been done trying to discover “bridge principles” between “is” and “ought,” the alethic-deontic status of these terms, and their social signification. Almost none of them, however, have examined the logical-semantic structure of deontic judgments in a manner that reveals their literal meaning. Although a non-descriptivist, it is Hare, surprisingly, who has provided us with perhaps the closest glimpse at such meaning, although, again, his analysis is completely disconnected from any substantive analysis of goodness. Hare’s very constrained views on imperative statements, however, ultimately limit the success of his account.

Hare tells us that a practical syllogism must contain at least one imperative premise. It follows from this, he asserts, that there can be no practical syllogisms comprised wholly of factual premises. From all of my previous analysis, it can be seen that we have significant grounds for doubting whether the former statement is true. As for the latter, it simply does not follow. Although Hare is careful to remind us of the existence of hypothetical imperatives, in which a practical judgment is held to follow from a statement of desire or want, which is clearly a statement of fact, his account encounters three fatal flaws: (1) he repeatedly conflates “indicative” and “factual;” (2) he does not seem to consider that hypothetical imperatives, or dare I even say categorical imperatives, ever comprise one or more of the premises of a practical syllogism; and (3) he fails to consider the possibility that imperatives could be derived from factual non-indicative premises.

Let us begin with the first error. The current classifications of grammatical phenomena are somewhat misleading. While indicatives are said to express statements of fact, it is quite wrong to say that indicatives are the only kinds of statements which do this. Conditional statements, for instance, are said to express what *would be*, yet is not an expression of what *would be* a statement of fact? To clarify, we can think of conditional statements as, themselves, descriptions of logical relationships between states of affairs, namely the relationships of sufficiency and necessity. These logical relationships are, indeed, factual. We might quibble over the characterization of “factual,” using it to designate only empirical, and not logical, states of affairs. Even so, this would not contradict my thesis, for although technically inductive assertions, conditional statements concerning empirical states of affairs can express probabilistic cause-and-effect relationships that are, in speech if not in *deductive* logic, represented using such conditional expressions as “If I hit you with this hammer, it will cause you to feel pain” or even such

² W. D. Hudson, *The Is-Ought Question: A Collection of Papers on the Central Problem in Moral Philosophy*, (London: Macmillan. 1969).

³ Gerhard Schurz, *The Is-Ought Problem: An Investigation in Philosophical Logic*, (Dordrecht: Kluwer Academic Publishers. 1997).

⁴ John R. Searle, “How to Derive ‘Ought’ from ‘Is,’” *The Philosophical Review*. 73, no. 1, (1964): 43–58.

⁵ Alan Gewirth, “The Is-Ought Problem Resolved,” *Proceedings and Addresses of the American Philosophical Association* 47, (1973): 34–61.

statements as “If you are about to drown, and nobody prevents you from drowning, you will drown.” Although grammatically, these are statements of what *would be*, statements of what *would be* are also statements of what *is*. In grammar, this distinction does have its place but, as we can see, it can wreak havoc with philosophy. If something is a factual statement, it does not follow that it is an indicative statement.

His second mistake concerns the assumption that neither hypothetical nor categorical imperatives can comprise the premises of practical syllogisms. While it is true that he considers hypothetical imperatives to be possibly constitutive of practical syllogisms themselves, he does not clearly consider them to be constitutive of premises *within* practical syllogisms. The further implied assumption, here, is that these premises would not be factual. “If I want to live, I must eat” is technically a factual assertion; it is intended to imply something along the lines of “Wanting to eat is necessary for wanting to live” or “My eating is a necessary condition of my living,” and Hare admits this. Could this statement not be paired with some other hypothetical or indicative premise, say, the indicative premise “I am on a desert island, and all there is to eat are coconuts?” Given these premises, it seems clear that we must conclude the following: “It is necessary for me to eat coconuts in order to live.” Hare might object and tell us that there must be an imperative statement here somewhere, or else the syllogism lacks prescriptive force. On the contrary, if we are to count wanting something as a sufficient reason for doing an action in question, which is illegitimate but which we will use since it is in any case what Hare’s own theory implies, then our syllogism does, indeed, have prescriptive force. Hareans will have to accept this if they accept the grounds of Hare’s theory. Here is where our analysis becomes interesting. If our conclusion, “It is necessary for me to eat coconuts in order to live,” seems somewhat superfluous, that is because it is; this is not the kind of thing we say in ordinary speech. In ordinary speech, we tend to “chop off” the clause “in order to live” and instead say simply “It is necessary for me to eat coconuts” or, as Hare suggests, “Eat coconuts.” In ordinary language, therefore, when we make statements about what it is necessary to do, we omit the sufficient condition or antecedent; we take it to be implicit. This, then, takes us to Hare’s third error, that of failing to consider that imperative statements can be, or even all are, derived from factual, non-indicative statements. On my account, imperatives are derived from conditionals. If, indeed, non-descriptivists were to assent to the claim that imperatives are derived from conditionals, then they would also have to admit that it is possible to derive a value judgment with prescriptive force from a factual statement, since imperatives, especially on Hare’s account, are said to be prescriptive or action-guiding.

Throughout his works, Hare impresses upon us that all “ought”-statements are derived from syllogisms containing imperatives. In fact, he offers an account of deontic terms as modal operators of necessity, much as I wish to. He states:

The ordinary modal ‘must’ does not have, like the deontic ‘must’, the property of prescriptivity; but it has an analogue to it. Most systems of modal logic contain either axioms or theorems to the effect that if something *must be* so, it *is* so. Now if, in deontic logic, we were to introduce a similar axiom, with the simple change of the proposition

governed by the operator into the imperative mood, we should have an axiom or theorem to the effect ‘If you must do something, then do it’, where the ‘must’ is deontic. The deontic ‘must’ is then treated in the same way as the ‘must’ of ordinary necessity, except that it governs an imperative. . .⁶

On Hare’s account, then, the antecedent of such statements as “If you must do something, then do it” is also usually implicit. As I have stated, however, it seems that deontic judgments that are not, themselves, imperatives are derived from conditional statements. Unfortunately, Hare never demonstrates the precise logical relationship between imperatives and “ought”- and “must”-judgments beyond his characterization of these latter two types of judgments as functions of choice and expressions of necessity, which is, in any case, indirectly related to imperatives. Let us, however, dig deeper. “Ought”- and “must”-judgments, it seems, are derived from conditionals. Hare, however, also seems to imply that imperative statements are derived from a conditional the antecedent of which is an “ought”- or “must”-judgment. If this is the case, which I think is apparent, then it follows that imperative statements are derived from “ought”- and “must”-judgments and not, as Hare seems to suggest, vice versa. The structure of his own theory seems to suggest this.

4.1.2 Actuality, Subjunctivity, and Conditionality

In order to place this thesis on firmer grounds and to further demonstrate that “ought”- and “must”-judgments are derived from conditional statements, let us dig even deeper. The thesis that deontic judgments are expressions of necessity is not new; it reaches back at least as far as Kant’s comments in the *Groundwork*. What ethical theorists have generally failed to demonstrate, however, is the exact logical and linguistic nature of such necessity. In the case of hypothetical imperatives, the nature of this necessity is clear enough, but in the case of categorical imperatives, this has classically seemed a bit more obscure. Gewirth, at least, has demonstrated that such a categorical imperative, also derived from a conditional, is possible. “Must” and “ought,” however, seem to have different logical and linguistic functions. Concerning “ought,” Hare tells us, “Its logical properties and function are closely analogous to those of the other modal operators like ‘it is necessary that’.”⁷ This, however, seems to be only partially accurate. If we interpret Hare’s comments as implying that “ought” and “it is necessary that” are basically the same, we run into some serious mistakes. As we shall see, if this were the case, then “ought” and “must” would be interchangeable. In applying this reasoning to our previous example, “If I want to live, it is necessary to eat coconuts,” it is clear, as I have previously stated, that we can, with full coherence, drop the antecedent and

⁶ Hare, *Moral Thinking*, 23.

⁷ Hare, *Sorting Out Ethics*, 136.

simply state “It is necessary (for me) to eat coconuts.” As Hare suggests, we can permissibly replace “it is necessary” with some alternative deontic expression. That “it is necessary” is translatable to “I must,” if not directly “I ought,” is clear if we examine the deontic patterns of other languages. In French, for instance, the phrase *Il faut que*, which is commonly translated as “It is necessary that,” is interchangeable with such phrases as *Je dois*, meaning “I must.” The same phrase is not, however, completely interchangeable with the phrase *Je devrais*, meaning “I ought” or “I should.” As in English, it seems, thus, that not just any translation or paraphrasing of “it is necessary” will do. This calls into question whether “ought” and “must” have the same logical function or meaning. Hare’s argument against the thesis that “ought” implies “can” is instructive of this:

Suppose that I say ‘I ought to go see him, but I can’t, because I don’t know where he is’. There need be nothing inconsistent in such a remark; and yet there would be if ‘ought’ always implied ‘can’. But there are, in fact, various ways in which ‘ought’ can be, as it were, weakened, so as no longer to possess the property which makes ‘ought’ and ‘cannot’ disagree.⁸

While “ought” might not always imply “can,” we must question whether this is applicable to “must” as well. It is evident, from common usage, that “must” usually, but perhaps not always, has a categorical sense in the way that “ought” does not. We can see this difference thus: If I must do something, it follows that I ought to do it, but if I ought to do something, it does not follow that I must. Certainly, there are odd cases in which “must” can become weakened and in which “ought” can become strengthened, but it seems to be the case that many of these rare cases of usage are, in one way or another, tied to a specific type of speech act. Thus, indeed, it seems that “must” and “ought” have these strong and weak tendencies, respectively. Both, however, seem to capture the sense of necessity inherent in conditional statements. It is, thus, our task to discover what logically or semantically distinguishes the two terms. In English, it is not as obvious just quite what characterizes this distinction. In romance languages, however, this difference is, perhaps, more perceptible due to the systematic and almost methodical pronunciation and spelling rules that clearly distinguish tenses, moods, and conjugations. Let us revisit the above distinction, in French, between *Je dois* (I must) and *Je devrais* (I ought, I should). The verb *devoir*, which is best translated as “to have to” or, more awkwardly but to the point, “to must,” is the infinitive basis of these two expressions. *Je dois* is clearly an expression of directive modality. While also usually considered an expression of directive modality, *Je devrais* is actually the conditional form of the statement *Je dois*. *Je devrais*, in fact, can also be used to say “I would have to,” which is telling in itself. This pattern is true of any French verb, whether it be *pouvoir* and *Je pourrais*, *vouloir* and *Je voudrais*, or *manger* and *Je mangerais*. It is through examining these linguistic patterns that we may also discover the nature of imperative statements. The French subjunctive forms of “to be” might help us. In French, if I want to say “Be nice!,” I would say *Soyez sympa!* *Soyez* is the subjunctive second-person plural

⁸ Hare, *Freedom and Reason*, 52.

conjugation of the verb *être*. This mood, in French, is acknowledged as an abbreviated expression of statements of desire, want, or necessity. This is easy to see if we examine such statements as *Je veux que vous soyez sympa!* (I want you to be nice!) and *Il faut que vous soyez sympa!* (It is necessary for you to be nice!). In English, we no longer say anything along the lines of “I want that you be nice” or “It is necessary that you be nice,” but the subjunctive mood is covertly at work here too. What is peculiar about imperative statements, in French as in English, is that we use the subjunctive form. We do not, for instance, say “Is nice!” when we want to command someone to be nice. It is through the conditional-subjunctive link that we might, then, see the distinction between “must” and “ought.” In French and other romance languages, again, this distinction is more evident. Consider such expressions as *Il faut que vous soyez sympa*. We see that the sense of categorical necessity contained within the statement “It is necessary for you to be nice” can easily be abbreviated to the imperative “Be nice” via the subjunctive mood. This is not clearly so with “ought”-judgments. This is where Hare’s comment that “ought” does not always imply “can” becomes especially relevant. As we shall see, it is, perhaps, more accurate to say that “ought” always implies “could.” To make this connection even clearer, we might, instead, say that “should” always implies “could.” If we understand both of these words as conditional terms, we will be able to understand the logical structure of “ought” and “should” as clearly distinct from that of “must.” Consider the statement *Vous devriez être sympa* (You ought to be nice). Recall that, in French, this very clearly takes the form of a conditional second-person plural conjugation of *devoir*. In other words, this statement also means “You would have to be nice.” Conditional statements, whether explicit or implicit, have a correlative subjunctive clause such as “if you were in situation X.” This, however, is not the same sort of subjunctivity as is implied by imperative statements derived from such statements as “It is necessary that X.” It is vital to note, then, that *Je devrais* and *Je dois*, while one is a conditional and one is not, do not share the same pattern of subjunctive correlativity. *Je dois*, for instance, necessitates no past-tense subjunctive if-clause. It is for the simple fact that *devoir* is not paired with the term *que* that a present-tense subjunctive is not correlative with it; consider, for instance, that *Il faut que* and *Il est nécessaire que*, both explicit assertions of necessity, are always paired with the term *que*. *Je dois* shares this same sense of necessity, yet neither a present- nor past-tense subjunctive if-clause is paired with it because the expression contains no *que* term, and this is because it always works on an infinitive. If it did contain a *que* term, that is, if the verbs that followed *Je dois* were not always infinitives, it would be paired with a present-tense subjunctive clause. Compare this with *Vous devriez être sympa* and all conditional statements in general; conditional statements entail if-clauses. Consider the statement *Je le ferai si je pouvais* (I would do it if I could, where “could” is past-tense subjunctive). If we simply said “I would do it” by itself, our listener might, depending on the context, be a bit confused. “If what?,” our listener might ask. “Under what conditions would you do it?” We do not, however, usually ask the same question when presented with the statement “I ought to do it” or “I should do it;” we do not, that is, usually feel the need to offer a correlative if-clause. This,

however, does not imply that these statements are not conditional ones. It is more probable, rather, that the antecedent is implicit within the given discursive context. The most common implicit antecedent, as has been previously stated, is “If I can (or could).” We begin to see, then, that it is the present-tense versus the past-tense subjunctive mood which partially distinguishes “must” from “ought.” It is important to note, however, that, in asserting that “ought”-statements entail past-tense subjunctive if-clauses, I am not implying that they always do. Instead I wish to express that, unlike “must”-statements, “ought”-statements alone are capable of past-tense subjunctivity, and this is precisely because they are incapable of expressing categorical obligation in the same way that “must”-statements are. We shall shortly see what I mean by this.

The question remains as to what characterizes these tenses and exactly how these characterizations contribute to the distinction between “ought” and “must.” We have established that both terms indicate necessity and that, at the very least, “must” is a deontic modal operator of necessity. The nuanced difference in their correlative subjunctivity indicates the equally nuanced difference in usage between the terms. Present-tense subjunctivity, we have said, is correlative with “must”-statements and their derivative imperative statements, while past-tense subjunctive clauses are correlative with “ought”-statements. We have also said that “must”-statements express a strong sense of categoricalness, while “ought”-statements do not usually express this same sense. I shall now attempt to explain whence derive these different senses. “Ought”-statements have the possibility of being weak, or otherwise not strongly categorical, because their implicit or explicit logical antecedents or consequents usually express a state of affairs that is not actual. Consider the statement “I ought to go to sleep.” The implicit antecedent “if I can,” a certain use of “can” that I will shortly elucidate, does not always express actuality. I might, for instance, be flying a plane. Although I recognize that I might be tired and that, therefore, I ought to go to sleep, I perceive that I ought to go to sleep if some other conditions are met that constitute a “can” for me. In other words, in uttering “I ought to go to sleep,” I might imply the clause “but I can’t because some condition X is not actual.” From this implicit “cannot”-judgment, and also taking into consideration our previous observation that “I ought” entails “I would have to,” we make the realization that our fulfilling the original “ought”-judgment requires a reformulation of the judgment to something along the lines of “I would have to do X if condition Y *were* met.” We should note, also, that it is not always an if-clause with which we are concerned with regards to “ought”-judgments; often, what concerns us is an only-if-clause. Thus, our statement changes to “I would have to go to sleep (*only if* some other conditions *were* met that constitute a ‘can’ for me).” This, however, has dramatic consequences, for our “ought”-judgment ceases to be the consequent and, instead, becomes the antecedent. “Ought”-judgments, it seems, can assume either sufficiency or necessity, inclusively. I might, thus, imply various types of assertions in my utterances of “ought”-statements. I might, for instance, utter a statement with any of the following senses: “If I *were* (*past-tense subjunctive*) at home, etc., then I would have to go to sleep,” “My being at home, etc. is sufficient for my having to go to sleep,” or “That I would have to go to sleep is

necessary for my being home, etc.” I might also utter an “ought”-judgment that is closer to something like “If I *had* (*past-tense subjunctive*) to go to sleep, then I would be at home” or “My being at home, etc. is necessary for my having to go to sleep.” The consequent of such conditional statements implies an if-clause, and such if-clauses, as we have seen, are past-tense subjunctive statements. Thus, if the “ought”-judgment is the antecedent, it is the grammatically subjunctive component of the logically conditional statement. If the “ought”-judgment, however, is the consequent, then it is the non-subjunctive, grammatically conditional component of the logically conditional statement. Since past-tense subjunctive statements express some non-actual, yet possible, conditions, we can see whence the revised dictum “‘Ought’ implies ‘could,’” as well as the weakness of “ought,” derives. Likewise, if the “ought”-judgment is the consequent, then the non-actuality of the statement becomes clear via the correlative past-tense subjunctive if-clause. One might object, at this point, by stating that I have said before that “must”-judgments are derived from conditional statements as well and that, thus, if we distinguish “ought”-judgments by their conditionality and sufficiency-necessity fluidity, there is no distinction here at all. We must, however, be attentive lest we become confused; we must not conflate conditionality and non-actuality. That “ought” implies “could” expresses both non-actuality and conditionality; it expresses non-actuality in the sense that “could” implies some state of affairs that is *possible* but not *necessarily* occurrent in the given circumstances in which I am uttering the “ought”-statement, and it expresses conditionality in the sense that it is part of a conditional statement. Conditionality and non-actuality, thus, must be recognized as two separate and distinct phenomena. With regards to the second concern, that of the sufficiency-necessity fluidity of “ought”-judgments, that “ought”-judgments have the ability to function as an antecedent or a consequent is vital, not fatal, to the distinction between “must” and “ought.” As we shall see, there is good reason to believe that, unlike “ought”-statements, “must”-statements are always consequents. It is here that the meaning of “must” becomes more apparent. Recall that I used the phrase “can for me.” I used this phrase in the previous steps of the argument in a purposely obfuscatory way, ironically, so as not to confuse the reader, and the necessity of this move will soon become apparent. Let us revisit the pilot’s remarks. He states “I ought to go to sleep.” He realizes, however, that he is flying a plane. From his perspective, thus, he realizes that this is something he *cannot* do. This, however, is not a “can” of possibility, for, surely, it is *possible* for him to sleep, although this would probably result in the plane’s crashing. This, rather, is a “can” of permissibility or, if you like, a “may.” Thus, we can at least begin to think of the difference between “ought” and “must” by considering them two different “layers” of judgment. First, the pilot makes the judgment that he ought to sleep. Second, he makes the judgment that he cannot do so because some other criterion that is necessary or sufficient, inclusive, for his having to sleep has not been fulfilled; this is how “ought” implies “I would have to.” But how does he determine this “can” or “may?” How, in other words, does he determine these necessary or sufficient (inclusive) conditions? Certainly, there are situations in which someone utters “I ought to do X” but in which she cannot, in the sense of possibility, do what

she judges she ought to do, simply because the necessary or sufficient conditions are not actual. In this situation, however, it is not merely the condition of non-actuality of ability which prevents the pilot’s decision to sleep. Rather, it is some additional form of *actuality* that is not present in “ought”-judgments themselves but in “must”-judgments. Our first hint that it is “must”-judgments that serve as the gatekeeper of such “ought”-judgments is that it was easy for this pilot, as it is often easy for many of us, to confuse a “cannot” of permissibility with a “cannot” of possibility. The pilot’s judgment, “I ought to sleep, but I cannot because I am flying the plane” is perfectly sensible and intelligible. If this is perfectly sensible and intelligible, then the following statement also must be: “I ought to sleep, but I must not because I am flying the plane.” “Must”-judgments, then, seem to express this strong sense of categoricalness because, while “ought”-judgments usually express some non-actual condition, “must”-judgments usually, if not always, express an actual condition that *overrides* the “ought”-judgment *by setting the “ought”-judgment’s “can” of permissibility or, in other words, its necessary or sufficient (inclusive) conditions*. Thus, we can see that, while Hare is correct, in one sense, to state that “ought” does not always imply “can,” he has overlooked another sense in which “ought” implies “can;” he and others have failed to note that “ought”-judgments that imply the “can” of possibility also imply a “can” of permissibility or, more simply, “must”-judgments. In other words, while “ought” always implies “could,” those “ought”-judgments which imply “can” are regulated by some underlying, standard-setting “must”-judgment. The objection might arise that the criterion of non-actuality that I have used does not serve my distinction well. “Must,” one might state, can be used in cases where the sufficient condition to which one refers is non-actual, for instance, in the statement “I must see him, but I cannot find him.” This, however, is a misinterpretation of what I mean by “non-actuality.” Certainly, both “ought”- and “must”-judgments can be meaningfully paired with some “can” of possibility. The “must”-judgments which are applicable here, however, reflect what I have been calling “weak” “must”-judgments, that is, ones which do not express categorical obligation. I cannot, for instance, be said to be obliged to do something it is impossible for me to do; I may want it very much or may wish that I could do it *if or only if* I were in a position to, but I can never be said to be categorically obliged to because the conditions which characterize this strong sense of “must” are non-actual, that is, they are not conditions which I can *possibly* fulfill. As we have seen, the criterion of non-actuality applies to situations whose possibility conditions are fulfilled but whose permissibility conditions are not fulfilled. This kind of non-actuality refers, thus, to the “can” of permissibility. If we think, once again, of “ought”- and “must”-judgments as “layered,” we can see that the “must” of categorical obligation is highly relevant for, although an “ought”- or weak “must”-judgment might be ineffectual due to the absence of the gate-keeping function of the “cannot” of possibility, there remains a “can” of permissibility which serves a further gate-keeping function by setting the standards which constitute the necessary or sufficient conditions of “ought”-judgments. Thinking of these, once again, as layered judgments, my point is not that “must”-judgments always perform this function but, rather, that they are the only ones that

"Ought"-judgments

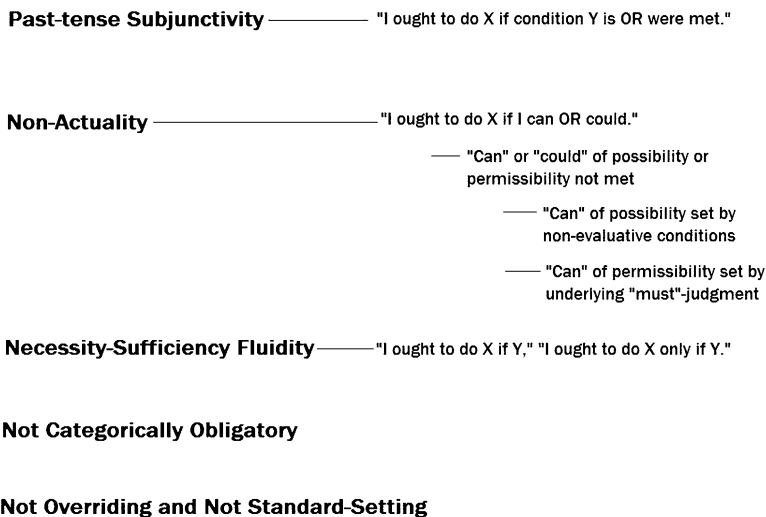


Fig. 4.1 Essential properties of 'ought'-judgments

can. Even strong "ought"-judgments, however we might characterize them, seem incapable of capturing the full sense of categorical obligation inherent in most "must"-judgments. Even if some "ought"-judgments could be shown to have this strength, however, this does not destroy the distinction, namely that there exist judgments of necessity that are separated by certain conditions of subjunctivity, necessity-sufficiency fluidity, non-actuality, strength in terms of categorical obligation, and overriding and standard-setting functions. Thus, I believe it is clear from this analysis that it is "must"-judgments which perform the last-mentioned function by acting as a gatekeeper for "ought"-judgments by setting their necessary or sufficient conditions. For a clear summary of these criteria and the relationships between them, see Figs. 4.1, 4.2, 4.3, and 4.4 below.

Another objection might arise concerning my characterization of "must"-statements as necessity-sufficiency rigid. Is it not true, one might ask, that I can intelligibly say "I must do X, *only if* Y?" The answer is yes; one, indeed, can intelligibly say this, and people do often say this. This does not mean, however, that "must"-statements have become sufficient, and not necessary conditions, although it is certainly possible that they *can* have become so. Remember that my point is not to demonstrate that "must"-statements cannot be weakened so as to become more like "ought"-judgments; my point, rather, is to demonstrate that, although they can become weakened, it is "must"-statements alone which are used to express categorical obligation. As expressions of categorical obligation, therefore, they express necessity and not sufficiency. Even, for instance, if we were to take the example "I must do X only if Y," we might be expressing the logical form " $M(i) \rightarrow Y(a)$,"

"Must"-judgments

Present-tense Subjunctivity ————— "I must do X if condition Y is met."

Actuality ————— "I must do X if I can, AND I can."
 — "Can" of possibility and permissibility both met

Necessity-Sufficiency Rigidity ————— "I must do X if Y; X is necessary for Y."
 — "Must" entails necessity only

Categorically Obligatory

Overriding and Standard-Setting

Fig. 4.2 Essential properties of ‘must’-judgments

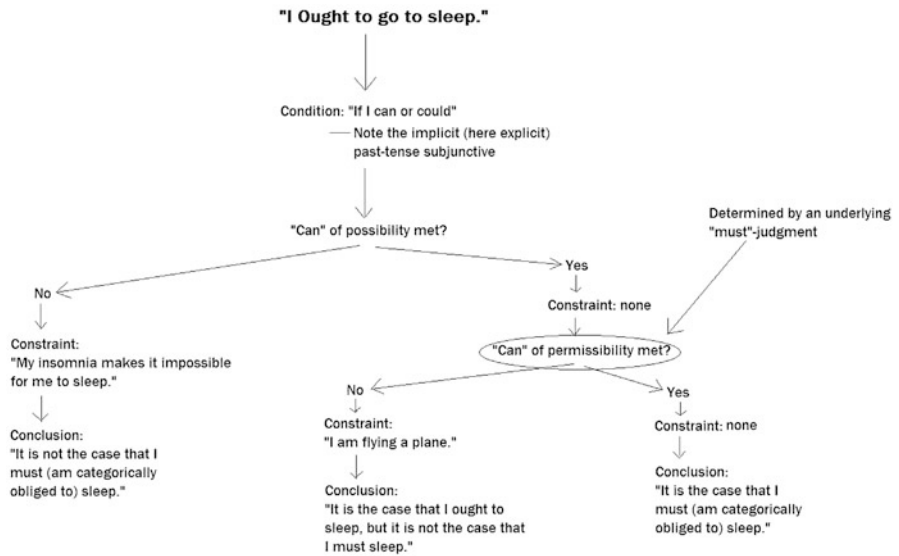
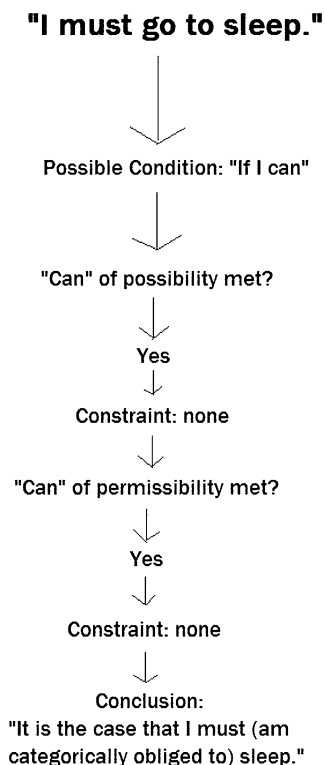


Fig. 4.3 Schematic representation of the logical-semantic regulatory process underlying decisionmaking

Fig. 4.4 Schematic representation of the logical-semantic process underlying 'must'-judgments



which would be the weakened form, but we also might be expressing the logical form $(Y(a) \rightarrow M(i)) \rightarrow (M(a))$ or “M(i) is necessary for Y(a), only if M(a).” In other words, even though one of the “must”-judgments is part of the antecedent of a larger conditional, the “only if” in the consequent, which expresses the necessary conditions of the entire statement, is still always a “must”-judgment. We can make this judgment as long and complex as we want; no matter how complex we make it, however, the statement’s only-if clause will always refer to some consequent which expresses necessity, and this necessity can only be expressed using a “must”-judgment. Even in the former statement “M(i) \rightarrow Y(a),” M(i) is the case *only if* Y(a). In other words, Y(a) is necessary for M(i) or, to clarify, if M(i), then Y(a) *must* be the case. Thus, even in the weakened form, it is not possible to have such a statement as “I must do X only if Y” without there being a “must”-judgment in the consequent.

Yet another objection might arise here, namely that categorical obligation does not follow from permissibility. How can we say, for instance, as in the above example, that it follows from my “ought”-judgment’s passage through the two “can” gates, that my sleeping is obligatory? Permissibility, however, is a flimsy concept; it is non-committal and indecisive. In my model, and I believe in the Gewirthian and other deontological models, there are two options open in every

choice: choice A is either obligatory or it is unjustifiable. Permissibility is, more than anything, a throwaway term referring to some set of choices from which it is not unjustifiable to choose. If this is all permissibility gives us, however, it either yields for us almost no prescriptive guidance or implicitly suggests to us that we are to completely disengage our rational decision-making faculties and simply choose at random, perhaps on the whim of some desire. This must be distinguished, however, from cases of what set theorists call incompleteness, in which one choice may be exactly as good as or at least indiscernibly as good as another; on my model, incompleteness must, thus, also be acknowledged as an inevitable possibility of social choice. On my model, however, there are two possible interpretations of categorical obligation: either (1) My given “ought”-judgment, after passing through the two “can” gates, entails a “must”-judgment; or (2) My “ought”-judgment is wrong or is otherwise a misjudgment. In the aforementioned model, therefore, if I utter “I ought to sleep,” I imply that I have a categorical obligation to sleep so long as my judgment passes through the two “can” gates. One might judge this to be extremely demanding of or onerous to the agent, but this is not the case. We must consider the psychological aspects of value as well. Freedom and its fulfillment, for instance, depend heavily on the psychological dispositions and general state of the agent. If I ask such a question as “Must I eat ice-cream at noon today?,” it might seem absurd to conclude that I would have some sort of categorical obligation either to eat it or not. In considering the moral relevance of psychological states, however, we must ask a vital question: “Having ruled out that my eating ice-cream at noon today will entail an immoral act (such as, perhaps, stealing it from someone), how will eating ice-cream at noon affect my psychological sense of well-being, and thus, freedom?” If it is the case that, all things considered, eating ice-cream at noon today will best serve the purpose of personal freedom fulfillment via some psychological mechanism (be it because the act is associated with a good childhood memory, because it somehow motivates you, because it helps you relax, etc.), then it is also the case that, according to the ideals of freedom fulfillment, agency enhancement, and individual excellence, you are categorically obligated to eat ice-cream at noon today even though it might not “feel” like an obligation. Simply stated, so long as there is no relevant conflict or dilemma between the advancement of your freedom via psychological fulfillment and some other salient criterion, it just is the case that you must do the thing which best advances your freedom via psychological fulfillment. Thus, we find that there exists, as some philosophers call it, “moral freedom,” but that this freedom exists, in fact, within the dictates of the sound moral system, and not, as it were, outside of or detached from it. Although we are not ever permitted, in a shallow sense, to do whatever we want, we are, indeed, permitted *ceteris paribus* to do whatever we want in the manner that I have described above. We will more clearly see why this is the case if we examine exactly how “must”-judgments set the necessary or sufficient conditions of “ought”-judgments. We can see that “must”-judgments are, in a very real sense, the final arbiter of decisions. But how do they achieve this? It is here that Gewirth’s theory helps us.

4.1.3 *Implications of Gewirthian Theory for Deontic Meaning*

Gewirth's theory emphasizes that there is a certain determinate relationship between "ought"-judgments, "must"-judgments and judgments of goodness. "Ought" and "must" are, in large part, linguistic tools for making choices. Choosing, it must be emphasized, is a process of inclusion and exclusion; one always chooses one thing rather than another, even if the choice in question is simply "I will do X rather than \neg X." The only way one can make such a choice, as Gewirth and Beyleveld have pointed out, is by making a value judgment concerning which choice it is that one ought to choose. In other words, whenever one makes a choice, it is always judged, using some set of values, to be a better or less worse choice than an alternative choice. Even in cases of *akrasia*, or weakness of the will, one judges, for instance, that "it is not so bad" to fulfill one's desire "just this once" or "every so often" *rather than* "doing the moral or virtuous thing." One might even judge one's choice to be a bad one, but, even so, one still chooses that that bad choice is better, or at least less worse, than some alternative choice; a bad choice must still be judged a better choice. In addition, it must be emphasized that there is no specific type of goodness one must cite to make one's choice; it need not be specifically intrinsic or extrinsic. The choice, for instance, might be a prudential one in which one judges "Choice X is better (for me) than choice Y." Choice is impossible without value judgment. We can see, therefore, that the judgment "X is better (or less worse) than Y" follows from the judgment "I ought to (or must) do X." We can take this a step further. If one makes the judgment "It is better (or less worse) (for me or otherwise) that I make choice A than that I make choice B" it would seem contradictory to make the simultaneous judgment "but it is not the case that I ought to make choice A." We must conclude, then, that positive "ought"- and "must"-judgments follow from such positive judgments of goodness concerning one's choices. To demonstrate this further, we might dress up our judgment of goodness as some general judgment of value: "It is more valuable (or less disvaluable) (to me or otherwise) for me to make choice A than choice B." If I judge my making choice A to be more valuable or less disvaluable than my making choice B, it could not possibly make sense for me to simultaneously utter, "but I ought not make choice A." Again, this judgment of goodness or value can be for any reason whatsoever; one might judge something good because it gives one pleasure, because it leads to a cure for cancer, or because it leads to mass genocides. It does not follow, of course, that one's judgment of goodness, even if prudential, is an accurate one. One might object using Jean-Paul Sartre's classic example of having to decide between joining the French resistance movement against Germany during World War II and staying with one's mother to look after and care for her.⁹ Let us suppose that one makes the decision to join the resistance movement. One, therefore, at some point, makes the

⁹ Jean-Paul Sartre, *Existentialism is a Humanism*, (New Haven: Yale University Press. 2007).

judgment “I ought to join the resistance movement instead of caring for my mother.” However, our objector might say, one really wants to help one’s mother at the same time and actually considers helping one’s mother more valuable or better than joining the resistance movement. He, therefore, chooses to join the resistance movement while simultaneously judging “It is better or more valuable for me to help my mother.” This, however, makes no sense, unless one makes such a judgment only after one has committed oneself to some inextricable situation within the resistance movement. This, however, is not what our objector is claiming. She is claiming, rather, that one can make the choice to join the resistance movement and judge helping one’s mother a better or more valuable decision *at the same time and in the same respect*. We must, however, be skeptical. One had to make a choice, and one had to choose between one (joining the resistance movement) or the other (taking care of one’s mother). It is at this point that we must seriously ask, if one is to make a choice at all between two courses of action, how it is possible, in the first place, for one to make such a choice. In other words, there has to be a psychological mechanism for choice if it is to be conscious, voluntary and not simply the result of a non-agentive impulse or reaction. What, then, is the psychological mechanism? As my account of rationality has shown, judgment is necessary and sufficient for all rational action. The psychological mechanism, then, must be some form of judgment. But then we must also remember that any rational action whatsoever necessitates a choice, whether it be the act of studying theoretical physics, snapping one’s fingers, or eating waffles. All choice, then, necessitates a judgment between choices. We must make a judgment of what is more worthwhile; in other words, we must make a judgment of value or goodness. It is from this initial judgment of value or goodness that agents are even able to make choices; without this initial judgment, all choice would be impossible. In other words, without this initial judgment, all “ought”- and “must”-judgments, with regards to choice, would be impossible. We see, then, the formation of a bi-conditional statement: “‘Ought’- or ‘must’-judgments concerning one’s choices are necessary and sufficient for statements of positive value or goodness concerning one’s choices.” Let us, however, not be confused by this. By this statement, I do not mean to say “‘Ought’- or ‘must’-judgments are necessary and sufficient for statements of positive value or goodness.” This is precisely why I qualified the antecedent and consequent with the phrase “concerning one’s choices.” Although we might admit that the judgments “A is good” and “A ought to exist” entail each other, these kinds of judgments are not, yet, those upon which I intend to focus. Keep in mind, however, that these statements do, indeed, entail each other. A further analysis of this, however, will be undertaken in Chap. 5.

Let us see how this affects the earlier objection that permissibility does not entail categorical obligation. The strangeness of this objection should, by now, be clearer. If we are to make a choice at all, we must have some process of final arbitration for determining which choice we ought to choose. Assuming that we have passed through the two “can” gates, and, thus, that we have determined that the action which our original “ought”-judgment endorses is permissible, we are left, as we have seen in this analysis, only with the standards of value or goodness which

regulate our “ought”-judgment; there are no “cannots,” and thus no “may nots,” which constrain us from making a final judgment, unless, of course, we judge that our standard is no longer or was never categorically obligatory. Furthermore, since no choice is possible without some judgment of goodness or, more precisely, some judgment of “betterness,” it can easily be seen that our making this final judgment necessitates a judgment of what is better or less worse. We must not, therefore, simply make a choice concerning what is permissible, for this is no choice at all; we still have not chosen a course of action. I hope it is clear, then, that the final judgment must be a “must”-judgment and that it is to be a judgment concerning what it is categorically obligatory for one to do, even if this categorical obligation is something as seemingly inconsequential as going to sleep at a certain time and under certain conditions instead of reading for an hour more. If I have chosen sleeping instead of reading, I have made a judgment that it is categorically obligatory for me to do so, at least in those specific circumstances.

The analysis I have undertaken here has important consequences both for deontological and virtue ethics. Some conceptual links between the two ethical models might already be apparent, but I will undertake a fuller examination of the various and intricate conceptual relationships between the two models in Chap. 5. For now, let us commit ourselves to developing even more tools for our analysis. Our next stop is to examine the significant implications of the current analysis on the concepts of rights and duties.

4.2 The Dialectical Structure of Rights and Duties

In this section, I argue that the concepts of rights and duties, so controversial in Gewirth’s theory, are nothing more than straightforward derivations from “must”- and “ought”-judgments.

From this, I conclude that such concepts as rights and duties have direct logical connections to aretaic or virtue concepts. Through this theoretical dissection, we begin to see clearly the conceptual bridge between dialectical deontology and virtue ethics.

4.2.1 *Relativism and the Correlativity of Rights, Duties, and “Must”*

While the existence of the concept of duty is relatively uncontroversial, that of the concept of rights is a contentious issue. Many arguments against the existence of rights, in general, and egalitarian-universalist rights in particular, have revolved around various charges of ethnocentrism and anachronism. Although horribly weak, moot, and thoroughly refuted, the general outline of these arguments against

the concept of rights will be briefly taken up here. It is important for me to take up such arguments in the context of my greater overarching argument because the specific counterarguments I want to offer, although along the same lines as many others who have written on rights, focus in on the specific relationship of rights and duties to “ought”- and “must”-statements. As both of these types of statements, as well as the concepts of rights and duties, are central to our discussion, it will be instructive to add to the refutation of these arguments by demonstrating the connections between these highly related concepts.

Let us, first, take up the charge of ethnocentrism. By any standard, this kind of objection seems like one from relativism or at least some kind of pseudo-relativism. There has been much criticism of relativism, and my counterarguments to it, while already implicit in my overall argument, will not even begin to do these counterarguments justice. However, I can certainly offer what I take to be the type of counterargument most relevant to my position. One main pseudo-relativist objection to the concept of right is that it is an oppressive, imposed standard of morality from Western culture upon other cultures. Let us see the logical mistakes and normative presuppositions that are present from the start in such arguments. The argument, we might immediately see, is implicitly contradictory since it already presupposes the existence of the rights of people of one culture not to be oppressed by people from a Western cultural background. The argument is self-defeating, for a specific universal standard has already been presupposed, namely “Do not impose your cultural standards on others because such imposition is, in some way, universally wrong.” To argue this point is but a few steps away from arguing for the existence of universal rights not to be oppressed. This is somewhat akin to the equally self-defeating subjectivist argument “X is good because I think it is good,” which states absolutely nothing of substance. It makes no sense to say “X is good because I think it is good” because there is no explanation of what one means by “good” and, if the subjectivist were to attempt to provide a definition for it, he would have to present us with a concept that is definable in terms of something other than his thinking something to be good, thereby contradicting his argument that X is good because he thinks it is good. Like the relativist and pseudo-relativist, he either contradicts himself or says nothing of substance; neither option is a good one.

Another objection might yet arise, namely that rights do not exist in other cultures. This, once again, seems to be a conflation of concepts and conceptions. Certainly, there are different conceptions of rights, mainly hierarchical-particularist and egalitarian-universalist ones. It is best to be prudent when any anthropologist or sociologist tells us that rights do not exist in other cultures. Needless to say, such anthropologists and sociologists are engaged in conceptual analysis where they probably should not be. This is a chronic problem in the sciences as well, in which oxytocin is sometimes referred to as “the love hormone” and in which morality is sometimes spoken of as residing in or as being constructed in specific parts of the brain. These are huge conceptual errors that have potentially grave consequences. The confusion concerning rights, again, seems to be a concept-conception conflation. At the heart of the concepts of rights and duties is the concept of practical

necessity. As I, Gewirth, and many others have at least implicitly argued, this concept of practical necessity is a “must;” in other words, the practical necessity at the heart of the concepts of rights and duties is a “must”-judgment, an expression of categorical obligation. When we reveal the essential conceptual framework of rights and duties, we see that there is no ethnocentrism here at all. That to which relativists and pseudo-relativists object is not this concept of rights, but rather the specific *conception* of rights that is egalitarian-universalist. Although we can also demonstrate, and Gewirth and Beyleveld do demonstrate, that egalitarian universalism is the morally acceptable framework of rights and duties, we can use the *concept* of rights to argue against these pseudo-relativists on their own grounds. Would such an objector refuse to acknowledge that Confucianism prescribes duties to parents and correlative rights of parents to their children’s respect (particularly with respect to filial piety) and assistance in old age? Is this relationship not an expression of categorical obligation? Do not various religions prescribe a certain categorical obligation to revere a god or gods, to treat people in a certain role-regulated way according to their gender, age, or hierarchical status? Are these not duties with correlative rights? Although there has classically been controversy over whether all duties and rights are correlative, Magnell has convincingly argued, through his distinction between conditioned and unconditioned rights and duties, that they are, indeed, correlative.¹⁰ Diverging from Magnell’s own line of argumentation, it is plain to see why this is the case. Consider the assertion “I must have freedom.” This statement has the important implication “Others must at least not interfere with my having freedom,” which in turn has two important implications. As I stated before, the concepts of rights and duties are based upon simple statements of practical categorical obligation or, in other words, “must”-statements; all practical “must”-statements imply assertions of rights and duties. If I utter “Others must at least not interfere with my having freedom,” then it logically follows that I imply “Others have a duty not to interfere with my freedom.” Once again, this relationship is clear in romance languages. In French, the former statement is translated as *Les autres doivent au moins ne pas empiéter sur ma possession de la liberté* and, in Italian, as *Gli altri dovrebbero almeno non ostacolare al mio avere di libertà*. The latter statement is translated as *Les autres ont un devoir de ne pas empiéter sur ma possession de la liberté* and, in Italian, as *Gli altri hanno un dovere di non ostacolare a la mia libertà*. The reader might have noticed that, in both cases, the words *devoir* and *dovere* serve as the basis for both “must” and “duty;” in their verb forms they mean “must,” while in their noun forms, they mean “duty.” Because these two words in English are the same words in French and Italian, a tautology is more readily apparent in such phrases as “I must do my duty” (*Je dois faire mon devoir, Devo fare il mio dovere*) and “A duty is that which one must do” (*Un devoir c’est ce qu’on doit faire, Un dovere è quello che si*

¹⁰Thomas Magnell, “The Correlativity of Rights and Duties,” *The Journal of Value Inquiry*, 45, no. 1, (2011): 1–12.

deve fare). The same or a similar pattern exists in Spanish, with *deber*, and in other languages as well.

With regards to the concept of rights, we can see a relevantly similar relationship. While “Others must at least not interfere with my having freedom” implies that others must refrain from depriving me of my freedom and, thus, that they have a duty to do so, it also follows from this statement, analytically, that I must *be granted* my freedom. When we formulate the expression in this way, therefore, we imply “I have a right to freedom” with the implicit clause “against all others.” What makes these expressions mutually implicit but seemingly different is the speaker we are addressing in our respective uses of “duty” and “right.” When referencing duties, we are referring to that which someone must grant to someone else, while, when referencing rights, we are referring to that which someone must be granted by others. In short, the duty-right distinction is a simple active-passive grammatical distinction; duties and rights are always and necessarily correlative. Representing this formally might make this relationship clearer:

$$\begin{aligned} & ((\text{HasRightToBeGranted}(x, y) \wedge \text{HasRightAgainst}(x, z)) \\ & \leftrightarrow (\text{HasDutyToGrant}(z, y) \wedge \text{HasDutyTo}(z, x))) \end{aligned}$$

entails

$$\begin{aligned} & ((\text{MustBeGranted}(x, y) \wedge \text{MustBeGrantedBy}(y, z)) \\ & \leftrightarrow (\text{MustGrant}(z, y) \wedge \text{MustBeGrantedTo}(y, x))) \end{aligned}$$

I hope it has become clear that rights and duties are not, and do not claim to be, mystical, mysterious social constructions that are based solely upon the whims of self-righteous oppressors; it is actually quite the contrary.

4.2.2 *Rights, Duties, and Anachronism*

A second kind of objection to the concept of right is the charge of anachronism, with which I will deal only briefly. The objector might, as Macintyre has, assert that the concept of rights has existed only since the Middle Ages in Europe, and then, only within the context of feudalism. This is not only a confusion of the concept-conception distinction, but it is also an implicit assertion that, because words change, concepts change. The first assertion, as we have seen, is plainly mistaken. Did conceptions of duties to the polis and, thus, rights of the polis against the individual, not exist in 400s BC Athens? Did and does Confucianism not establish duties to family and society and, thus, rights against the individual to fulfill these

duties? Did the Iroquois Confederacy, as well as a number of native American peoples, not rest on a system of gift-giving and reciprocity that, itself, was founded on certain categorical obligations of chiefs or leaders to their respective peoples? Certainly, none of these social configurations prescribe specifically egalitarian-universalist rights and duties, but they prescribe rights and duties nonetheless. Indeed, any society that has held any conception of owing has thus also held particular conceptions of rights and duties. What would be more accurate to state on the part of anthropologists, sociologists, and historians is that the Western egalitarian-universalist *conception* of right is, perhaps, new but that the concept of right is not new at all. In fact, I believe a true, full conception of agent egalitarian universalism is extremely new and, furthermore, that it is held by very few people. What do I mean by a true, full conception of agent egalitarian-universalist rights? I am talking about a position that takes as equal the rights of people of all races, all gender identities, all sexualities, all classes, all cultural backgrounds, all religions, those with disabilities, mental illnesses, various body figures, and all other accidental features that a person may have. As long as the individual is an agent, that individual has rights equal to all other agents. This is a position I believe very few people today hold, and, yes, it is very new, but its newness is completely irrelevant to its soundness.

4.2.3 Concluding Remarks

In this chapter, we have explored the distinction between “ought”- and “must”-judgments, their relationship to judgments of goodness, and their relationship to the concepts of rights and duties. In our first discussion, we found that “ought” and “must” are distinguished by their respective prescriptive weakness and strength and that this difference is the result of various distinguishing criteria which demonstrate “ought”’s ultimate dependence on “must” and the standards upon which this “must” is based. In our second discussion, we found that some judgment of positive goodness concerning one’s choices is both necessary and sufficient for all “ought”- and “must”-judgments concerning one’s choices. A positive judgment of goodness concerning one’s choices is thus necessary and sufficient for any choice one makes. In our third discussion, we found that rights and duties embody the same sort of necessity inherent in practical “must”-judgments and, thus, that right- and duty-claims always imply some correlative “must”-judgment that expresses practical categorical necessity. As we are beginning to see, it is this concept of categorical necessity that undergirds and serves as a basis for all “must”-judgments, “ought”-judgments via their underlying “must”-judgments, right- and duty-claims, and judgments of goodness. Without the concept of categorical necessity that underlies all of these sorts of judgments, judgment itself, let alone practical or theoretical judgment, would be impossible. Now that we have

undertaken an at least somewhat sufficient analysis of these relevant concepts, let us move on to the next chapter, in which we will survey and analyze various salient positions with respect to a sound characterization of goodness and value. We will ultimately synthesize the present analysis, as well as other insights from dialectical theory, with these models to form a far-reaching aretaic and assertoric theory of value.

Chapter 5

Virtue and the Search for Intrinsic Goodness

In this chapter, I use the conclusions concerning the conceptual inextricability of deontic and aretaic concepts from Chap. 4 and other insights from dialectical theory in Chap. 3 and begin to apply these analyses to the concept of rationality I characterized and defined in Chap. 2. The goal of this process, to be attained in Chap. 6, is to demonstrate the existence of a linguistic-conceptual framework in the concept of rationality. By acknowledging and analyzing such a framework, I boldly argue, we can finally define “good.” This is, I claim, because goodness and value find their meaning only within what I call the Rational Framework. I begin in earnest to advance toward this goal in this chapter by taking up some epistemological challenges concerning value intrinsicity and by exploring some promising theoretical frameworks.

5.1 Magnell’s Challenge

In this section, I take up a challenge Thomas Magnell presents concerning value intrinsicity: What kinds of criteria make a property, in general, an intrinsic one, and how do we apply this to values and goodness, if at all? His criteria, I argue, provide a definitive method for testing whether value is or is not intrinsic. I take his criteria with me into further sections and apply it more thoroughly in the next chapter.

Toward the end of his “Evaluations as Assessments, Part II: Distinguishing Assertions and Instancing Good of a Kind,” Magnell concludes that, although usually used attributively, theorists have not clearly demonstrated that “good” is not sometimes used predicatively. Although not truly a formal challenge, his is an implicit call to all naturalist or, otherwise, descriptivist theorists to seek out a sound basis for intrinsic goodness. This is a challenge I will take up here. First, however, let us analyze Magnell’s argument and see how he has come to his conclusion.

5.1.1 *Probing the Predicative-Attributive Distinction*

After an analysis of such property distinctions as that of essence and accident, relational and non-relational, and supervenient and non-supervenient, Magnell concludes, on the grounds that all such properties can be further distinguished as intrinsic and extrinsic, that the intrinsic-extrinsic distinction is that which is most relevant to discourses on goodness; it does not follow from this, of course, that the other aforementioned distinctions are irrelevant. Furthermore, as he demonstrates, there is much in terms of analysis that we can draw from the intrinsic-extrinsic distinction. As I have pointed out elsewhere, Magnell has elucidated a central problem in non-descriptivist and even some descriptivist theses, namely that their adherents have classically assumed that all evaluations are assessments or that evaluative assessments constitute “peculiarities” that do not apply to assessments in general.¹ As a result of such assumptions, the possibility that statements of goodness are sometimes factual has gone overlooked or has remained under-analyzed. This kind of thinking, he shows, is flawed.

Magnell claims that evaluative assessments are “fundamentally comparative” and, as such, instantiate extrinsic properties. Such properties are signified by dependent signifiers or terms whose true or false application to an individual given in an assertion is dependent upon their true or false application to other individuals to which the term is applied. Intrinsic properties do not share this feature. Instead, they are signified by independent signifiers, or terms whose true or false application to an individual given in an assertion is not thus dependent. The examples which Magnell offers of each are “wide” and “wooden,” respectively. Dependent signifiers such as “wide,” therefore, are dependent upon comparison to other wide things. Wooden things, on the other hand, are wooden whether or not they are compared in a certain manner to other wooden things. The objection might arise here that the very notion of dependent signifiers suffers from an infinite regress, for how, one might ask, can something be wide only if other things are also wide, for would that not imply that those other things must also be wide only if other things are also wide, and so on *ad infinitum*? The objection, in other words, is that a term whose meaningful application depends on its comparative function in relation to other terms which likewise depend is incoherent. Magnell seems to anticipate this kind of objection by highlighting the following important point: an extrinsic property of an individual always depends on the intrinsic property to which it corresponds. The extrinsic property of being wide, therefore, is dependent on the individual’s intrinsic property of width. Thus, whether or not an intrinsic or extrinsic property is instantiated, there is always an at least implicit instantiation of some intrinsic property.

Magnell also analyzes the classifications of adjectives by their respective roles and work within assertions. He stresses that semantic roles are regarded *per se* by virtue of adjectives’ ability to occasionally change roles. An adjective’s semantic

¹ Magnell, “Evaluations as Assessments, Part II.”

role can either be logically predicative or logically attributive and intensionally closed and extensionally open or intensionally open and extensionally indeterminate, respectively, although logically attributive adjectives can sometimes be intensionally closed and extensionally determinate. Following Peter Geach, he classifies term "A" in the phrase "AB" as a logically predicative adjective if the predication "is an AB" splits up logically into a pair of predications. What this means is that, in order for "A" to be logically predicative, both "A" and "B" must meaningfully apply to the subject as individual terms. An example Magnell gives of semantic predication is "This is an American university." The term "American" is predicative because it is possible to separate the terms logically and have them apply meaningfully to the subject "this," as in the assertions "This is American" and "This is a university." Semantic attribution, however, follows a different pattern. An adjective is semantically attributive when "A" and "B" do not split up logically due to "A"'s dependence on "B" as a class of individuals to which "A" refers. In order for us to make a meaningful assertion with an attributive adjective, we must refer to a class of things via comparison. His example, "This is a tall dwarf," is instructive of this. In this assertion, "tall" is an attributive adjective because it refers only to the class of dwarves; the logical separation of the two terms, therefore, results in the incomplete predicate "is tall." The difference, in this respect, between predicative adjectives and attributive adjectives is that predicative adjectives refer to the same class of individuals both explicitly and implicitly, whereas attributive adjectives refer to the subclass explicitly and the superclass implicitly. In the example "This is a tall dwarf," "tall" is attributive because we refer to the subclass "dwarf" explicitly and to the superclass "person" implicitly. It would, thus, be non-sensical to describe a dwarf as tall without this clarification.

With regards to the semantic work of adjectives, Magnell stresses that adjectives work *with* rather than *on* nouns terms. What this means is that adjectives provide a clearer reference or set of references for noun terms rather than altering them. Adjectives, Magnell explains, work with nouns in two ways. Adjectives may work extensionally in that they designate a specific class of things by combining the two separate extensions or references to classes of a noun term and an adjective term. The assertion "This is a wooden desk" is an example in which the adjective "wooden" works extensionally. It works in this manner because "wooden" is a term used to refer to the class of all wooden things and "desk" is a term used to refer to the class of all desks. Adjectives may also work intensionally. Adjectives work in this manner when the adjective's and noun term's respective intensions combine to synthesize an extension. The adjective "forged" works intensionally in the assertion "This is a forged Rembrandt" because it refers to the subclass "forged Rembrandt." The term "Rembrandt," however, is not used intensionally because it does not refer to the subclass "forged Rembrandt." Although a Rembrandt is a type of painting, it belongs to a class of paintings specifically created by the artist which cannot, therefore, be forged. The assertion, thus, contains an intensional adjective *per se* because it refers to a subclass to which its noun term does not refer. As with an adjective's semantic role, its semantic work can also differ. Adjectives can work extensionally or intensionally depending on the context in which they are used.

Magnell elucidates a link between semantic role and semantic work. The work of an adjective, he asserts, determines its role. When an adjective works extensionally, it plays a semantically predicative role. This means that, in the assertion “This is a forged painting,” “forged” and “painting” are both used extensionally and can be logically separated to refer meaningfully to the subject “this.” By contrast, when an adjective works intensionally, it plays a semantically attributive role. Such adjectives as “tall” and “wide,” therefore, have indeterminate extensions and open intensions and cannot be logically separated to apply meaningfully to the subject. We see, thus, the distinct difference between such statements as “This is a forged Rembrandt” and “This is a forged painting.”

In elucidating the differences between intrinsic and extrinsic properties, types of property signifiers, and the role and work of adjectives, Magnell is able to precisely and more accurately characterize the distinction between claims and assessments. These two types of assertions, he claims, are characterized by “classifications of properties, of terms signifying properties, and of adjectives.” In a claim, an intrinsic property is instanced, the signifier of the property is independent, and the signifier is also intensionally closed and extensionally determinate. In an assessment, however, an extrinsic property is instanced, the signifier of the property is dependent, and the signifier is also intensionally open and extensionally indeterminate. Magnell, thus, concludes that when we make an assertion with a simple indicative sentence using a term which signifies a property, the term and the property it signifies must have all the characteristics of either claims or assertions.

5.1.2 *The Definitive Criteria for Value Intrinsicity*

As one might see, Magnell’s argument is directly relevant to the adjective “good” and the property of goodness. In distinguishing the two classes of assertions, Magnell paves the way for discussion on evaluations patently instancing *good of a kind*. Magnell investigates whether sentences in which “good” is used as an adjective are used to make assessments or claims. He begins, as he does with other examples of assertions, by examining its semantic role. “Good,” he explains, plays a semantically attributive role in many cases. This is clearly evident from the term’s inability to be logically separated from the noun term with which it works. He provides some examples in which this is the case. “This is a good knife” is one such sentence in which, if separated, “good” cannot stand on its own, but must refer only to the term “knife” without explicit reference to its subclass. Such a subclass could conceivably be that of hunting knives. In such a case, the phrase “good knife” does not have the same meaning as “good hunting knife.” In this way, we see that “good” is an dependent signifier, that is to say it is inherently comparative. It is from here that we might perceive the main issue more clearly: although such uses of “good” are semantically attributive, we cannot rule out that “good” is semantically predicative in some cases. As Magnell

states: "We have seen that a word can play a semantically attributive role in some sentences even as it plays a semantically predicative role in others." It would be "too hasty," therefore, to make the "broad generalization" that all evaluations that employ "good" as adjectives are assessments. As food for thought, Magnell lays out G.H. von Wright's three classes of goodness. The first, instrumental goodness, is the capacity of something to serve some purpose with which it is characteristically associated, technical goodness pertains to the ability of someone to perform to a certain capacity, and utilitarian goodness pertains to the potential or ability to achieve certain ends or to be useful in some respect. Von Wright, he explains, does not demonstrate how the three classes are related, although he does state that it is possible that they are related by a generic kind of goodness. This goodness, it seems, is generically characterized by functionality, although this generic characterization is certainly inadequate by itself. If functionality, alone, is inadequate to yield for us a discovery and accurate characterization of intrinsic goodness, then which criteria can help us narrow our search? Here, too, Magnell provides us with a hint. According to his analysis, he concludes that, in order for "good" to have a semantically predicative role, we would need to find "a predicate with an extension that necessarily includes as a subclass of its extension the extension of the predicate of the given sentence." In the assertion "This is a good knife," for instance, this is clearly not the case, as the assertion of goodness might not also apply to the assertion "this is a good butter-knife" or "this is a good steak knife." Magnell, thus, does indeed offer us a challenge: discover a usage of "good" that conforms to these criteria, and you will have discovered intrinsic goodness. Magnell, I believe, has taken the right approach here. The first step toward the discovery of anything is to understand, at least in some outlined form, the criteria which accurately characterize that which one is attempting to discover. Since Magnell has done such a fine job of elucidating these criteria with regards to intrinsic goodness, we may now attempt to approach its discovery. We must, however, proceed with caution, for there are many pitfalls we must be careful to avoid. These pitfalls include possible mischaracterizations of functionality, as well as a failure to appreciate the role of comparison in statements of goodness that Magnell has laid out. If Magnell is correct, however, it implies that such thinkers as Judith Jarvis Thompson are incorrect in holding that "there is no such thing as being just good," and that those who hold, for instance, Michael Zimmerman's position are correct.² Since the failure to incorporate both of these concepts into previous accounts of goodness has undoubtedly contributed to what Magnell calls the "stale" "old line of thinking" and since these concepts are integral to our understanding of goodness in general, we will need to grapple with some epistemological issues first.

²Thompson, "The Right and the Good." *Journal of Philosophy*. 94, no. 6 (1997): 135; M. J. Zimmerman, "Defending the Concept of Intrinsic Value," *Recent Work on Intrinsic Value*. Ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman. (Dordrecht: Springer. 2005): 153–168.

5.2 Problems in Searle's Epistemology of Function

In this section, I explore some epistemological problems concerning the relationship between function and value in John Searle's epistemological theory. I explain that, contrary to Searle, function is not value-dependent. It is necessary to prove that function is value-independent, I explain, because function is a concept that is central to goodness and value more generally. A failure to take up Searle's argument would leave my later arguments open to objections, although perhaps weak ones, concerning their circularity or presuppositions of value or normativity.

By far, John Searle has set forth one of the best-argued positions for the value-dependent nature of function. This view, that function is a product of value judgment, is one against which I shall argue. Function, as I will later further explore, is a concept that is fundamental to goodness. If Searle's thesis that function is value-dependent is correct, then it follows that the argument I shall later present on value in general rests on a fundamental presupposition of value. This supposed presupposition of value in naturalistic arguments is one which Hare, Hampton, and many others have criticized as circular, self-defeating, or otherwise unjustifiably presumptuous. My aim in this chapter is to show that my arguments do not fall prey to such presuppositions of value. To demonstrate this, it will be necessary to argue against the claim that function is value-dependent. If the reader already agrees with me that function is value-independent, then she or he might as well skip ahead to the next section.

Searle has been a monumental figure in ethical naturalism, and his numerous writings have served well in defending the legitimacy and soundness of its central premises. It is vital to recognize, however, that he takes a very different approach to naturalism than do such virtue theorists as Foot, Anscombe, Geach, and Slote. This difference hinges on Searle's implicit acceptance of constructionism and dismissal of essentialism. As we shall see, it is this denial of essentialism which leads him to certain self-defeating epistemological conclusions. In his work, *The Construction of Social Reality*, one of his most pervasive claims is that value judgment and function are two intricately interwoven phenomena and, more importantly, that function is the product of value judgment. Unfortunately, Searle does not do much analysis of value judgment itself, besides his analysis of rights, duties, and obligations in general and their supposed social creation and construction. It is, perhaps, because he does so little analysis of the structure of value judgment itself that he is led to the erroneous conclusion that function is the product of value judgment.

5.2.1 *The Social Construction of Function*

"The complex ontology," he states, "seems simple," while the "simple ontology seems difficult." This, he claims, is because "social reality is created by us for our purposes and seems readily intelligible to us as those purposes themselves" and,

thus, because “once there is no function, no answer to the question, What’s it for? we are left with a harder intellectual task of identifying things in terms of their intrinsic features without reference to our interests, purposes, and goals.” Such a claim seems simple enough if we consider his comments that “Cars are for driving; dollars for earning, spending, and saving; bathtubs for taking a bath.”³ These functions, he states with some credibility, are invented or constructed ones. Such functions, it is also easy to see, are constructed with reference to our values. Even these functions, however, although they are observer-relative and, thus, ontologically subjective, are still usually epistemologically objective. Even if money is something that is socially constructed, for instance, it is absurd to hold up a dollar bill and state “This is not money, since money is socially constructed. Money isn’t real.” All of this is invalid inference. Money, indeed, is real; the mere fact that something is socially constructed does not, as Searle rightly acknowledges, exclude it from being real. We are not considering, of course, the possibility that dollars are no longer in use as money and, thus, that they have ceased to be money. In such a case, we might rightly hold up a dollar bill and state “This is not money.” This kind of social construction is not very controversial; it is, in fact, quite straightforward. The inferences he draws from these claims, however, are untenable.

Once one analyzes the basic premises of his argument, it is not difficult to see where Searle goes wrong. Let us, first, take his following comments on non-man-made phenomena:

As far as our normal experiences of the inanimate parts of the world are concerned, we do not experience things *as* material objects, much less as collections of molecules. Rather, we experience a world of chairs and tables, houses and cars, lecture halls, pictures, streets, gardens, houses, and so forth. Now all the terms I have just used involve criteria of assessment that are internal to the phenomena in question under these descriptions, but not internal to the entities under the description ‘material object.’ *Even natural phenomena, such as rivers and trees, can be assigned functions, and thus assessed as good or bad, depending on what functions we choose to assign to them and how well they serve those functions. This is the feature of intentionality I am calling ‘the assignment—or imposition—of function.’*⁴ (italics mine)

At first, it seems as though all Searle is implying is that, in addition to the function that rivers and trees actually have, we construct other functions on top of these functions. In other words, we might think that he is implying that our impositions of function are limited only to judging, for instance, that a particular river is good for diving into because it is deep and there are no predators or that a particular stream is good for cooling oneself; we might call these a “diving river” and a “cooling stream,” for instance. This, however, is exactly not all Searle is implying. He makes his point explicit: “The important thing to see at this point is that functions are never intrinsic to the physics of any phenomenon but are assigned from outside by conscious observers and users. *Functions, in short, are never*

³ John R. Searle, *The Construction of Social Reality*, (New York: The Free Press. 1995): 4.

⁴ *Ibid.*, 14.

intrinsic but are always observer relative."⁵ How does Searle get away with making such a far-fetched claim? The answer is simple: he ignores the essence-accident distinction and the relevance of formal and final causality. In order to see why this claim is a far-fetched one, let us ask a simple question: "What makes identification possible?" In other words, how is it that we even begin to identify objects in the world as objects in the first place? How is it that we look at a cat and say "this is a cat" and we look at a mat and say "that is a mat?" As I will demonstrate, if we take Searle's position, we will have to admit that it is simply our values which make the distinction between what is a cat and what is a mat; there is not really a distinction; it is all in our heads. Simply on our intuition, there seems to be something very wrong with this. Intuition, of course, is not argument, so we will have to demonstrate, logically, why this is problematic.

Let us, again, take up the first question I posed, namely "What makes identification possible?" Putting aside the obvious neurophysiological mechanisms which make identification possible, let us discuss the concepts relevant to identification. Identification, we know from the law of identity, is that $x = x$; x can also be a , b , c , or d , but it can never cease to be x . In our everyday lives, we use (or at least cite) the law of identity; we cite it when we choose which clothes to wear, which books to pack in our bag, and we cite it when we need to tell a door apart from, say, an alligator. But we have still not answered the question: how is it that we come to the conclusion that a given thing that we observe, whether through sight, smell, taste, touch, etc., is, indeed, the thing we observe? How do we come to the conclusion that a door is a door and not an alligator or that our books are books and not windmills? Put simply, we interact with these things, and we discover how they interact with us. In other words, we need to find out what it is that they *do* and how it is that they *behave*. Is it not plainly evident from the differences inherent in how a door versus an alligator interacts with us that they *function* differently and, furthermore, that they have *intrinsically different* functions? Certainly, I can assign values to these functions; for the alligator, I might say that it functions badly with respect to me, whereas a door generally facilitates, or functions well, with respect to me. This, however, proves nothing. Different things are nonidentical because they have intrinsically different functions *independent of value judgments*. We can demonstrate this using a general, perhaps even golden, rule: Identification by function, in some respect, is always prior to imposition of function. Imposition of function, we must admit, needs *something* to be imposed upon. We must identify something *first* if we are to impose some function upon it, and we can only identify something by ascertaining its function. We can see this clearly if we consider that value judgments are propositional attitudes, that is that they are statements which always concern *something*. If I say "I believe," I must have an object of my belief, and if I say "I desire," I must have an object of my desire. Likewise, if I say "I value" or "X is valuable," I am saying that I value *something* and that *something* is valuable. *But I cannot value anything if there is not something to value in the first place that really*

⁵ Ibid.

exists and which is distinguishable via some observation of that thing's intrinsic behavior or, rather, its function. We must ask, then, how it is logically possible for value to precede function and for value to render function and not vice versa. This, as we might observe, is the classic mistake of existentialism that, somehow, it makes some sort of sense to state that existence precedes essence. Even if we consider the case of a hammer, which one might admit is a socially constructed object with a socially constructed function, we must identify the object, first, not necessarily as a hammer, but as *something* that behaves in a certain way; it is made of metal or plastic, it is oblong, it is small or large, etc. This is something which Searle's own premises admit, including "The sheer existence of the physical object in front of me does not depend on any attitudes we may take toward it" and "[An object] has many features that are intrinsic in the sense that they do not depend on any attitudes of observers or users. For example, it has a certain mass and a certain chemical composition."⁶ This, however, is indicative of intrinsic, value-independent function. Metals, for instance, can be more or less conductive, reactive, etc. The metal's essence, therefore, is those properties that the metal has which make it that metal. We can apply this system of classification to anything in nature, such as trees, rivers, rocks, and clouds, and all without referencing values or having their respective *teloi* in some way determined by agent values.

5.2.2 *The Problem of Classification in the Absence of Essence*

Another example that Searle gives seems to challenge this view, but, if we are careful, we can see the logical mistakes in this example as well. He states:

We are situating these facts relative to a system of values that we hold. It is intrinsic to us that we hold these values, but the attribution of these values to nature independent of us is observer relative. Even when we *discover* a function in nature, as when we discovered the function of the heart, the discovery consists in the discovery of the causal processes together with the assignment of a teleology to those causal processes. This is shown by the fact that a whole vocabulary of success and failure is now appropriate that is not appropriate to simple brute facts of nature. Thus we can speak of 'malfunction,' 'heart disease,' and better or worse hearts.⁷

This, however, is question-begging, for how can we meaningfully refer to something called a heart if we have not, first, identified it via its intrinsic function, that is, via a function which it has independently of our values? Here is where the essence-accident distinction is positively vital. Suffice it to say, we cannot simply go around classifying things in any old way we choose; we might well use different words than we now use, but words change nothing about function. We might be

⁶ *Ibid.*, 10.

⁷ *Ibid.*, 15.

tempted to assign the heart the function, as Searle suggests, of oxidation, and we might, thus, define the heart as “some lump of matter which tends to oxidize.” The essence and *telos* of hearts, however, is intrinsic. We can demonstrate this by asking a simple question: “Are all things that are lumps of matter which oxidize hearts?” If so, then we will have to call a massive amount of things that exist in the universe hearts. This, however, might not bother the constructionist. He might tell us, for example, “Fine, everything that oxidizes is a heart; it doesn’t matter how we use words.” While it actually does matter how we use words, and that is a different discussion altogether, word usage is actually irrelevant here. It is, rather, observation that is important. I might utter the word “heart” and be very dissatisfied that the word “heart” refers to lots of things in the universe; this was not what I wanted to say. I wanted to say that this specific thing, here, in the body, which pumps blood, is that which I wanted to denote; I did not want to denote a multitude of things in the universe. So be it, then; let us satisfy the constructionist, for the sake of argument, and accept that the word “heart” denotes a multitude of things. The point, however, is that I did not want to denote *all* of these things. I, instead, wanted to denote something that is *one* of the multitude of things which oxidizes. I choose, then, to make up a new word to refer to the thing to which I want to refer, say, “ Φ ”.

How, then, can I describe a Φ ? To answer this, I need to narrow down how a Φ , specifically, behaves. I realize that my prior definition, “that which oxidizes,” is too expansive; there is a multitude of things in the universe that oxidize. Φ s are, however, a *subclass* of things which oxidize, which is its own class of things. We try, then, to narrow this down further. What other things can oxidize? Organic materials and inorganic materials can oxidize; these are subclasses of things which can oxidize, and, although Φ s are also a subclass of things which can oxidize, we still have not narrowed down a precise definition and classification of Φ s. Well, we would say, “ Φ ” is certainly a subclass of “organic materials,” but what other subclasses does “ Φ ” belong to? “Things that are made of lipids, proteins, and carbohydrates” is a subclass of “organic materials.” “Organ” is a subclass of “things made of lipids, proteins, and carbohydrates,” and “ Φ ” is a subclass of “organ.” This is a very over-simplistic classification, but it gets the point across. Φ s, as a subclass of organs, must be identified by some specific characteristics it has in order for it to be distinguished as a particular type of organ. It is at this point that we must ask, “What does a Φ do that no other organ does?” Among numerous other things, it pumps blood, of course. Contrary to what Searle and other constructionists would like to claim, therefore, this is the precise reason why (1) we need not make any value judgments about the class of things which we are describing; and (2) we cannot assign any function we please to anything we please. Certainly, a heart is something which oxidizes, but it does not follow from this that all things which oxidize are hearts; this is a basic logical mistake. A heart, therefore, no matter how much we would like to, cannot be defined merely as something which oxidizes, since this only vacuously describes what it does and fails to distinguish between the various other things in the universe which also oxidize. As I stated before, we can change our word usage regarding the object in question, but this changes nothing about the object in question’s intrinsic function.

These points about the obligatory non-arbitrariness of classification hold, also, for Searle's other comments concerning the classification of "heart." He states:

Thus given that we already accept that for organisms there is a value in survival and reproduction, and that for a species there is a value in survival and reproduction, and that for a species there is a value in continued existence, we can *discover* that the function of the heart is to pump blood, the function of the vestibular ocular reflex is to stabilize the retinal image, and so on. When we discover such a natural function, there are no natural facts discovered beyond the causal facts. Part of what the vocabulary of 'functions' adds to the vocabulary of 'causes' is a set of values (including purposes and teleology generally). It is because we take it for granted in biology that life and survival are values that we can discover that the function of the heart is to pump blood. If we thought the most important value in the world was to glorify God by making thumping noises, then the function of the heart would be to make a thumping noise, and the noisier heart would be the better heart. If we valued death and extinction above all, then we would say that a function of cancer is to speed death. The function of aging would be to hasten death, and the function of natural selection would be extinction.⁸

Once again, Searle misses the point. As explained above, we cannot randomly assign whichever functions to whichever phenomena we please. Many things make thumping noises; are they all hearts? Many things speed death; are they all cancer or aging? What this points to is a special logical relationship between something and its function, namely "Function X belongs exclusively to some class Y, if and only if function X does not belong to any other class Z; $Y \neq Z$." We can think, for instance, of rivers: "Flowing belongs exclusively to the class of rivers, if and only if flowing does not belong to any other class." We know that rivers flow, but so do oil spills, blood, and a number of other things. We can take an even more extreme example to bring the point home. If we follow Searle's logic, we would have to admit that, if the thing we valued above all else was splashing around to glorify God, we would have to accept that the function of rivers is to splash around and glorify God. This, as we have seen, is a mistake. There are, for instance, pools, oceans, creeks, ponds, puddles, and particular rivers to splash around in for the purpose of glorifying God. If we follow Searle's logic, however, we will have to accept that the function of all of these different objects are the same. We must, however, remember that, in order to identify anything in the first place, in other words to identify its essence, we need to understand how it functions; we identify and define things by how they function or behave. If this is the case, however, and Searle is implying that all of these different things have the same function, then he is claiming that they have the same essence. He is, in other words, claiming that different things are identical, which is a violation of the law of identity. It is clear, rather, that all of the aforementioned bodies of water, as well as things that are not water, can have a function imposed upon them consisting in splashing around to glorify God. It does not follow from this, however, that this is their intrinsic function and intrinsic *telos*. Rivers have specifically different tendencies than lakes, creeks, oceans, or pools. Rivers do not merely flow; they have taxonomical

⁸ *Ibid.*, 15.

features which distinguish them from the aforementioned bodies of water. Although we do often make value judgments in our classification of things in nature, we need not make such value judgments. How, for instance, is the function of nebulae or plasma derivable from agent values? Remember the “golden” rule stated above: Identification before evaluation. One might object that evaluation is a form of identification, and one would be correct. This, however, is no objection. Instead, what it reveals is the necessary logical relationship that, in order to identify something as “a good X,” it is a necessary condition of such identification that we, first, identify the thing as “an X.” Evaluation is a subclass of identification, but it is a species of identification which has, as its necessary condition, the condition that a functionality judgment be made. Essence precedes existence, not vice versa.

The reader might be beginning to notice something interesting, although certainly not original, about function or *ergon*, namely that something X’s *ergon* is a necessary condition of X’s being a good X. This is significant. To see why, let us refer to the previous analyses of necessity and “must”-judgments. We established earlier that making a positive value judgment concerning one’s choices is necessary and sufficient for making an “ought”- or “must”-judgment. If this is the case, then any assertion of necessity indicative of the categorical obligation of a voluntary action entails an assertion of positive goodness. Let us state, for instance, that voluntary action X is necessary for or obligatory for Y. If what we have said above is the case, then “X is necessary for or obligatory for Y” entails “X is better for or good for Y.” This is also significant, since it means that it is possible to judge some function to be good and, furthermore, that it is possible to judge some function to be better than some other function for some specific *telos*. In our final section, the importance of this will become apparent.

Other flaws in Searle’s epistemology of function are important to point out for the purposes of my argument as well. Let us take, for instance, the following dilemma which he proposes:

What is so functional about functions, so defined? Either ‘function’ is defined in terms of causes, in which case there is nothing intrinsically functional about functions, they are just causes like any others. Or functions are defined in terms of the furtherance of a set of values we hold—life, survival, reproduction, health—in which case they are observer relative.⁹

Essentialists have never doubted, however, that functions are defined in terms of causes, namely material cause, formal cause, efficient cause, and final cause. What I and most essentialists mean by “function” is not some mysterious set of other-worldly causes. Intrinsic function, put simply, is that which a thing does essentially or which things of a certain class do essentially, that is, that which a thing does which nothing else does or that which things of a certain class do which things of no other class do. In Korsgaard’s terms, it is “how [an entity] does what it does;” it is the function found via an examination of the entity’s Aristotelian form.¹⁰ Nozick

⁹ *Ibid.*, 16.

¹⁰ Christine M. Korsgaard, “Aristotle’s Function Argument,” *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Ed. Christine M. Korsgaard. (Oxford: Oxford University Press. 2008).

characterizes function in different terms, namely "Z is a function of X when Z is an effect of X and X was designed or shaped (or maintained) to have this effect Z. Such designing or shaping can be done either by a human designer or by the evolutionary process. In either case, the designing itself seems to be a homeostatic process whose goal is that X produce Z."¹¹ While this characterization strictly concerns the concept of tendency or probability of an object's acting in a certain way and not, at least explicitly, the concept of *telos* (which is what moderns might describe as a normative concept), it is wholly compatible with my view, which is really Korsgaard's interpretation of Aristotle's view, which is purposely far more general or generic than Nozick's. My definitional purpose here is to argue that, in whatever way we conceive of something's function as its "doing" or its activity, Searle's assertion that there is "nothing intrinsically functional about functions" is logically absurd. If a function is a function, then it is a function, and if something is intrinsically itself, then a function is intrinsically a function.

5.2.3 Functions, Mental States, and the Background

The most intriguing and, at once, the most problematic feature of his account of functionality is his account of mental states. It is intriguing because it exposes reality as value-filled; it is problematic because it seems to covertly contradict Searle's own thesis on function. He states:

Because mental states, both conscious and unconscious, are themselves intrinsic features of the world, it is not strictly speaking correct to say that the way to discover the intrinsic features of the world is to subtract all the mental states from it. We need to reformulate our explanation of the distinction to account for this exception as follows: Intrinsic features of reality are those that exist independently of all mental states, except for mental states themselves, which are also intrinsic features of reality.¹²

As intrinsic features of reality, according to Searle, our minds are inherently imbued with value judgment, value judgment which we impose upon the world in the form of functionality judgments. The problem is that, in order to make such an argument that there are intrinsic features of minds, one must also implicitly argue that there are intrinsic features of minds *whereby they function*. If, therefore, Searle wishes to argue in this way, then he must admit that *there is at least one intrinsic function that exists*, namely the intrinsic function of minds which consists in constructing functions from values. If all mentally constructed functions are to be constructed from values, then does this not imply that the function of the mind *must not* be a mentally constructed function? In other words, in order to argue that all functions are mentally constructed, one must concurrently argue that the mind functions in a certain way that it does not, itself, determine, invent, or create. It seems, thus, that

¹¹ Nozick, *The Nature of Rationality*, 118.

¹² Searle, *The Construction of Social Reality*, 11.

one must concurrently argue that the mind functions (1) in a way that is mentally constructed and (2) in a certain way that it does not, itself, determine. If mental construction is a determinate act, which it is, then (1) and (2) entail a contradiction. We can see this contradiction more clearly in Searle's discussion of what he calls "the Background." He states:

In my writings on issues in the philosophy of mind and the philosophy of language, I have argued for what I call the thesis of the Background: Intentional states function only given a set of Background capacities that do not themselves consist in intentional phenomena. Thus, for example, beliefs, desires, and rules only determine conditions of satisfaction—truth conditions for beliefs, fulfillment conditions for desires, etc.—given a set of capacities that do not themselves consist in intentional phenomena. I have thus defined the concept of the 'Background' as the set of nonintentional or preintentional capacities that enable intentional states of function.¹³

Searle claims that the function of the Background consists in "facilitating linguistic interpretation and... in facilitating perceptual interpretation, [which] are extended to consciousness generally..."¹⁴ Are we to believe that the function of the Background, that which enables and is necessary for all imposition of function in the first place, is a product of value judgment? If so, how is this logically possible? Even if one were to object that it is not logically impossible to state this, which might technically be correct if we consider infinite regress logically possible, one would have to argue, as I stated, on grounds of infinite regress. Let me demonstrate why this is the case. If all function is value dependent, then Searle would have to admit that the function of the Background, and thus the Background itself, is value-constructed. The problem, however, is that the Background serves as the basis for all judgments of function. We would have to admit, therefore, an infinite regress and, simultaneously, that there is no basis for value- and function-construction at all. We would have to, therefore, make the following claim: "Functions are value-constructed from the Background which is value-constructed from the Background which is value-constructed from the Background which is value-constructed from the Background... *ad infinitum*." I greatly hesitate to grant that this is a logical possibility. If it is not a logical possibility, and there is, indeed, a Background, then there must be at least one value-independent function.

5.2.4 Concluding Remarks

When Macintyre commented on the failures of the Enlightenment in *After Virtue*, I believe he missed the mark a bit. While I think the Enlightenment yielded both philosophical failures and successes, he and I seem to diverge on what, exactly, belongs in the "success" and "failure" categories. While Macintyre holds that

¹³ Ibid., 129.

¹⁴ Ibid., 133.

deontology and utilitarianism, and ultimately emotivism and intuitionism, belong in the "failure" category and that nothing belongs in the "success" category, I take a different perspective. Deontology and rationalism, I believe, belong in the "success" category, even though Kant's and others' attempts have yielded greatly flawed theoretical frameworks. For all his theoretical flaws, however, Kant's work was seminal to the great work that is being done in deontology today. In the "failure" category, however, I am with virtue theorists and essentialists in highlighting the fatal mistake of such philosophers as Bacon and Descartes in abandoning essence and formal and final causality. It is this abandonment of Aristotelian metaphysics, I believe, that has led to such constructionist mistakes as we find in Searle's theory.

The symptoms of Aristotelian abandonment are endemic, however, in all constructionist thinking. Constructionism rightly emphasizes the constructed nature of social reality and the constructed nature of conceptions and even mental states. To be aware of social and cultural construction is of utmost importance, for instance, in being self-aware and in understanding the emotions and psychological states of others. To understand that things exist which are socially constructed is, of course, also of vital importance for understanding value. It follows from none of this, however, that all or even most things in social reality are socially constructed. In other words, it is possible to take the constructionist thesis too far. It is possible, for instance, to state that ethics is socially constructed, that disease is socially constructed, and that such entities as logic and mathematics are socially constructed. It is important, however, to be very attentive to what I am actually saying here. I am not saying, for instance, that ethics has not been constructed in different ways in different cultures or at different times, or that gender or race have not been constructed in the same way, or that disability has not been so constructed. What I am saying, instead, is that such categories *themselves* are not always socially constructed. Certainly, ethics is not *itself* entirely socially constructed, although it can be constructed, for instance, in an egalitarian or particularist manner. Many have made the case that gender, race, and disability are *themselves* socially constructed, and this may well be the case. I would like to argue, however, that it is more likely than not that such social phenomena as gender, for instance, additionally have some basis in essence. Does that essence have an automatic basis in sexual organs or hormonal systems? I would hazard to say that this is actually not the case. Might this essence have its basis in some particular performance or behavior? Perhaps. The important thing to recognize is that, if something is socially constructed, then it is socially constructed around *something*. Might that something be arbitrary, harmful, or just plain absurd? Certainly. But is that something unreal? Not necessarily. If something is constructed in a certain way, it does not automatically imply that that something *itself* is socially constructed, and this is an important distinction. It has been said, for instance, that AIDS is a social construction, but it would be more accurate to state that AIDS has undergone social construction. The disease, itself, exists non-socially, but the social stigma associated with it and the subculture that has developed around it are constructed; but this does not mean that the disease, itself, is constructed. This is not to state, also, that a transgender woman is not a woman. If we

were to demonstrate that a transgender woman were not a woman, we would, first, have to define “woman,” would we not? In other words, would we not have to find the essence of womanhood, whether it, itself, is a social construction or not? How do we find such an essence? We ask everyone who uses the word “woman,” of course; and I emphasize *everyone who uses it*. “To what are you referring when you use the word ‘woman?’” we would ask. Out of common usage, we will usually find a common, although perhaps very general and vague, concept by picking out the common features of all given definitions. That is not all. We would have to look at the images in societies where such a word is used and examine the history of those cultures in which the word is used. If we do all of this and closely examine the usage of the word “woman,” we will find its essence, whether that essence has been constructed or not. At the end of the day, I think we will find that a transgender woman is a woman, that a transgender man is a man, and that a person with an alternative gender or no gender essentially has that alternative gender and is essentially genderless, respectively.

“What,” you might ask, “is this talk about a constructed essence? Is this not contradictory?” Not at all. When we construct houses, do we not construct them out of different materials? Do they not come in all shapes and sizes? Of course they do. These houses, however, are still houses; yet they are also constructed. Let us move this into the realm of the conceptual. Rationality, for instance, can be said to be constructed in various ways: I could have it be game-theoretical or expected-utility theoretical, Kantian or Humean, classical or romantic, etc. Are not all of these constructions, however, constructed from *something*? As the house is still a house whatever accidental form it takes, is rationality not still rationality whatever form it takes and however we conceptualize it? It is a fault of essentialists and constructionists alike that they implicitly, and often explicitly, assert that essentialism and constructionism are incompatible when, in reality, I cannot conceive of how they could be incompatible. If the conception is a construction, then the concept is its essence, and if the essence is constructed, then it is constructed from something else that has a non-constructed essence. Otherwise, what constructionists would have us believe is that it is possible to construct something from nothing, and this I cannot accept.

Now that we have examined the relationship between *ergon* and value, namely that value is dependent upon *ergon* and not vice versa, we are ready to investigate the concepts of comparison, necessity, and *ergon* in relation to goodness as a natural property. To provide the last piece of the puzzle before our full-scale analysis of this relationship, we turn, now, to the virtue theory of Philippa Foot.

5.3 The Life Framework: The Significance of Foot’s Virtue Theory

In this section, I explore the thought of virtue theorist, Philippa Foot, and her arguments on goodness and the “grammar of goodness.” Her arguments, I claim, bring certain vital points concerning the semantic application of goodness to light,

but her argument leaves other vital components missing. I use this as a jumping-off point for my next and final chapter, where I demonstrate, using Carnap's classic insight concerning linguistic-empirical and linguistic-conceptual frameworks, that what I call the Rational Framework, as a transcendental framework of value, solves this problem.

5.3.1 *Function, Necessity, and the Grammar of Goodness*

Thus far, we have seen that there are at least three concepts that are central to goodness. As nearly all virtue theorists admit, function or *ergon* is one such central concept. *Ergon*, alone, however, explains nothing. That is why we elucidated another concept that is central to goodness, namely that of necessity. Necessity, we see, is the central concept of deontology. Although it is not the case that anything that is good for something else is necessary for that thing, we will see, in the next section, why necessity is still central in such "good-for"-judgments. Magnell showed us that comparison is also a concept that is central to goodness. Comparison, however, is not, strictly speaking, central to *goodness*. It is, rather, central to what I will be calling "betterness," or that property of goodness that is instantiated by comparison and not by goodness itself. In other words, the concept of goodness, itself, does not elucidate the metric whereby we can compare good things to each other; only "betterness" can do this. We will eventually see why this way of speaking is necessary. For now, let us more superficially explore the relationship between function, necessity, and comparison. As was implicitly acknowledged above, we can begin to see a close relationship between function and necessity. Some function X, for instance, can rightly be said to be good for function Y if function X is a function that is necessary for function Y. This, however, is simply the tip of the iceberg; much more will be said about this later. My main point here is that *ergon*, a central aretaic concept, and necessity, a central deontic concept are intricately intertwined in all judgments of goodness, intrinsic or not. Having stripped away much overlying theory, we can see, thus, where virtue ethics and deontology converge. Such a relationship is most apparent, and actually implicit, in the virtue theory of Philippa Foot. In her work, *Natural Goodness*, she implicitly, and often explicitly, acknowledges such a relationship, although she does not acknowledge it as a convergence of virtue ethics and deontology. Once again, virtue theorists and deontologists tend to have the visceral tendency to stay as far away from each other as possible, and perhaps there is some reason to be found in this from the historical trends of ethical theory. The relationship, however, is clear. Citing Anscombe, Foot states:

Anscombe is pointing . . . to what she has elsewhere called an 'Aristotelian necessity': that which is necessary because and in so far as good hangs on it. We invoke the same idea when we say that it is necessary for plants to have water, for birds to build nests, for wolves to hunt in packs, and for lionesses to teach their cubs to kill. These 'Aristotelian necessities' depend on what the particular species of plants and animals need, on their natural habitat, and the ways of making out that are in their repertoire. These things together determine

what it is for members of a particular species to be as they should be, and to do that which they should do. And for all the enormous differences between the life of humans and that of plants or animals, we can see that human defects and excellences are similarly related to what human beings are and what they do.¹⁵

Here, Foot not only emphasizes the relationship between necessity, function, and goodness, but she also emphasizes, here as extensively elsewhere, that “moral judgement of human actions and dispositions is one example of a genre of evaluation itself actually characterized by *the fact that its objects are living things*.”¹⁶ (italics mine) This is, perhaps, the most ingenious part of Foot’s argument, for, although she does not explicitly acknowledge it, she is invoking the concept, first explicitly set forth by Carnap, of empirical-linguistic frameworks.¹⁷ Let me explain. A linguistic-empirical framework, on Carnap’s account, is a specific basis for which or paradigm within which linguistic-empirical concepts, and only linguistic-empirical concepts, exist. If, therefore, we are to have an ontology that we use as a reference for all of our claims concerning what exists, we must use concepts relevant to this framework. Thus, Chomsky’s famous example of “Colorless green dreams sleep furiously” is not literally meaningful because there is no linguistic-empirical framework within which green things are colorless, in which dreams, themselves, sleep, and so on. It is arguable that the linguistic-empirical framework was something to which Wittgenstein alluded in his criticism of Kant’s presupposition that there exists a noumenal, in addition to a phenomenal, reality. The linguistic-empirical framework is essentially a conceptual framework; in order to be meaningful, all of our concepts must fall within this framework and not within the realm, for instance, of square circles. Searle’s concept of the Background, indeed, is such a type of framework. How, however, does Foot’s theory invoke the concept of linguistic-empirical frameworks? Life itself, she seems to suggest, constitutes its own conceptual framework. As she states:

Judgements of goodness and badness have, it seems, a special ‘grammar’ when the subject belongs to a living thing, whether plant, animal, or human being. . . I think that this special category of goodness is easily overlooked; perhaps because we make so many evaluations of other kinds, as when we assess non-living things in the natural world, such as soil or weather, or again assess artefacts either made by humans as are houses and bridges, or made by animals as are the nests of birds or beavers’ dams. But the goodness predicated in these latter cases, like goodness predicated to living things when they are evaluated in a relationship to members of species other than their own, is what I would like to call secondary goodness. It is in this derivative way that we can speak of the goodness of, for example, soil or weather, as such things are related to plants, to animals, or to us. And we also ascribe this secondary goodness to living things, as, for instance, to specimens of plants that grow as we want them to grow, or to horses that carry us as we want to be carried, while artefacts are often named and evaluated by the need or interest that they chiefly serve. By contrast, ‘natural’ goodness, as I define it, which is attributable only to living things

¹⁵ Foot, *Natural Goodness*, 15.

¹⁶ *Ibid.*, 4.

¹⁷ Rudolph Carnap, “Empiricism, Semantics, Ontology,” *Revue Internationale de Philosophie* 4 (1950): 20–40.

themselves and to their parts, characteristics, and operations, is intrinsic or 'autonomous' goodness in that it depends directly on the relation of an individual to the 'life form' of its species.¹⁸

Her invocation of a "special grammar" related to life is a clear instantiation of a framework, what I will call the Life Framework. The Life Framework is defined as the paradigm within which all concepts directly or indirectly concerning life can be meaningfully applied. Her implicit contention is that the language of natural or intrinsic goodness is only applicable to concepts within this framework and is only meaningful within this framework. As she states: "My belief is that for all the differences that there are, as we shall see, between the evaluation of plants and animals and their parts and characteristics on the one hand, and the moral evaluation of humans on the other, we shall find that these evaluations share a basic logical structure and status. I want to suggest that moral defect is a form of natural defect not as different as is generally supposed from defect in sub-rational living things."¹⁹ This, I believe, is truly a break with twentieth-century ethics, for, although other virtue theorists have hinted at this, Foot has most clearly enunciated the logic of the position. I believe, however, that the position is even further enunciated through explicit reference to empirical-linguistic frameworks. Just as green things are not meaningful outside of what might be called the Color Framework, or the framework containing those concepts pertaining to the class of things that can have color, intrinsic goodness, Foot claims, is not meaningful outside of the Life Framework, or the framework containing those concepts pertaining to the class of things that can have life and, thus, goodness.

Foot narrows down the precise pattern of usage within the framework. In order for something to have natural goodness, it must "play a part in the life" of an organism. She explains:

Consider, for instance, a sentence such as 'The blue tit has a round blue patch on its head.' This is superficially like 'The male peacock has a brightly coloured tail', but in a way of course it is not. For, on the assumption that colour of head plays no part in the life of the blue tit, it is in this respect quite unlike the colour of the male peacock's tail: there would be nothing wrong with the blue tit in my garden *in that* it had a drab-coloured head; and the peculiarity might or might not accompany a defect. So how are we to distinguish these two types of propositions? Or how, again, are we to distinguish the case of leaves rustling when it is windy from that of flowers opening when the sun comes out? It is natural to say that the rustling of its leaves plays no part in the life of a tree, whereas pollination is gained by a display of scent and colours in sunshine.²⁰

But what specific criteria, she rightly asks, constitute something's "playing a part" or "being necessary" in the life of an organism? The answer, she contends, comes through our consideration of the link between Aristotelian categoricals and "teleology in living things." She states:

¹⁸ Ibid., 26.

¹⁹ Ibid.

²⁰ Ibid., 30.

I should say that in plants and non-human animals these things all have to do, directly or indirectly, with self-maintenance, as by defence and the obtaining of nourishment, or with the reproduction of the individual, as by the building of nests. This is ‘the life’ characteristic of the kind of animal with which the categoricals here have to do. What ‘plays a part’ in this life is that which is causally and teleologically related to it, as putting out roots is related to obtaining nourishment, and attracting insects is related to reproduction in plants.

She explains that, in order to understand what makes an organism a good organism of that class of organisms, we must know the nature of that organism’s life cycle; in other words, we must know its *to ti en einai*, *ergon*, and *telos*. If natural goodness, in general, is “measured” by the standard of life itself, then the entire teleology of goodness is constituted by life and continued life or survival. That is why Foot concludes that “The way an individual *should be* is determined by what is needed for development, self-maintenance, and reproduction: in most species involving defence, and in some the rearing of the young.”²¹ As she states:

We could say, therefore, that part of what distinguishes an Aristotelian categorical from a mere statistical proposition about some or most or all the members of a kind of living thing is the fact that it relates to the teleology of the species. It speaks, directly or indirectly, about the way life functions such as eating and growing and defending itself come about in a species of a certain conformation, belonging in a certain kind of habitat. This is why the noise made by the rustling of leaves is irrelevant in this context while the development of roots is not.

The teleology, thus, is determined by the framework itself and by the specific context in which function and necessity are employed together within this framework. Thus, different organisms have different essences, and, thus, on Foot’s account, different specific *teloi* determined by their specific essences, as all final causes are. While these *teloi* might be essentially different due to the organisms’ different essences, they are all members of the superclass of things that have the *telos* of life and survival. This *telos* is not the result of mere tendencies of these organisms’ *ergoi*. Instead, this *telos* is the result of “good”’s sole meaningful application within the Life Framework.

5.3.2 *The Magnitude of the Implications*

Foot’s achievements are momentous, and for several key reasons. The main reason is that she has given a more determinate structure to goodness. She has demonstrated how it is used, in what contexts it is used, and how we are to measure it using a clear standard. Furthermore, she has demonstrated how *to ti en einai*, *ergon*, *telos*, and necessity are fundamental to the logical structure of judgments of goodness. In some very real sense, she has demystified goodness as an unnameable, indescribable, or intuitive concept and has demonstrated, by discovering its framework of usage and application, how goodness is actually grounded upon an

²¹ *Ibid.*, 33.

empirical-linguistic basis. Many elements of goodness, however, remain shrouded. We know in what framework “good” is used and how it is used within that framework, but what does “good” actually mean? Foot has provided a firm basis for concluding that function and necessity play a central role in regulating the word’s use, but she has not provided a definition of it, a definition which would most likely even further elucidate our understanding of its use. In addition, she has highlighted how some sort of comparison between good, better, and best is possible, but the applicability of the sort of comparison she has highlighted seems extremely limited. To demonstrate this limitation, let us take the following scenario. You are walking in a jungle, and, suddenly, you are faced with a hungry lion. You face a choice: somehow escape from the lion and, thus, have the lion die of hunger, or submit to the lion but sacrifice your life. Which, on Foot’s theory, is the better option? Certainly, if you escape from the lion, and it dies from hunger as a result, that is good for you; but it is quite bad for the lion. Alternatively, if you choose to submit yourself to the lion, thus dying but ensuring that the lion does not die of hunger, that is quite good for the lion and quite bad for you. This is not, however, what we wanted to know; we wanted to know what was intrinsically better and not what was better for which party. In other words, Foot has demonstrated a basis for comparison of natural goodness via “good-for”-judgments, but she has not demonstrated a basis for comparison that is not confined to mere reference to particular individuals or parties. The thesis that nothing can be good but can only be good-for is one that has been persistent at least since Wittgenstein’s comments on the subject and which, in recent years, has been forcefully upheld by Thompson.²² Since this thesis has been so persistent, it will be important to demonstrate how it can be successfully rebutted although, importantly, not refuted. Certainly, function and necessity will play a role in demonstrating how something can be meaningfully called good and not simply good-for, but we will have to elucidate the relationship between these concepts a bit more. As it turns out, Wittgenstein is technically correct to state that all statements of goodness are “good-for”-statements, but this, as we shall see, does not discount the concept of intrinsic goodness. Can something, for instance, be good for goodness? If so, we will, first, have to be careful that we are actually expressing something and not simply asking, and then answering, a vacuous question. As we shall see, this question is definitely not a vacuous one; indeed, it reveals the nature of intrinsic goodness and not simply goodness-for.

Let us ask another vital question: If there is a Goodness Framework, perhaps comprised of various sub-frameworks, then, based on what we have already said about the salience of the concept of necessity, would the existence of this framework, itself, not be intrinsically good? Would the Life Framework, and perhaps even life itself, not be intrinsically good? Foot is correct to imply that the Life Framework is that empirical-linguistic framework within which goodness is meaningful, but I do not believe that she is correct to state that this is the only framework

²² Ludwig Wittgenstein, “A Lecture on Ethics,” *The Philosophical Review*, 74, (1965): 3–12; Thompson, “The Right and the Good.”

within which it is meaningful. If, for instance, an Existence Framework, which should be self-explanatory, is necessary for the Life Framework, then, if the Life Framework is intrinsically good, then the Existence Framework would be intrinsically good as well, since such a framework is necessary for the Life Framework. If the reader is skeptical, however, that is justified at this point. We have not, as of yet, demonstrated how or why the necessity of the Existence Framework would render the Existence Framework intrinsically good. Furthermore, we have not demonstrated that intrinsic goodness exists, in the first place, beyond the Life Framework. I ask the reader, however, to hold her skepticism until she has read the final section of this chapter. In the final section, I will elucidate the exact relationship between function and necessity, and I will demonstrate how another concept, that of tendency, is vital for our understanding of goodness as well. From these concepts, I will present a naturalistic definition of goodness. From there, I will examine “good-for”-judgments and will show that discovering intrinsic goodness, by itself, is inadequate to provide a basis for hierarchical judgment, of what is better or best. Due to “good”’s inadequacy to yield a “better,” I will establish a concept which I call betterness, which I believe to be necessary above and beyond goodness alone. As we shall see, we will need to cite more frameworks. As I have already mentioned, the Existence Framework will be important, and so will Foot’s Life Framework, but a central framework will become apparent—that of the Rational Framework. As I will show, it is only through analysis of the Rational Framework that we can discover the answer to the questions: “What is necessary for betterness?” and, also to a large extent, “What is necessary for goodness?” Goodness, it will become evident, is much easier to define than betterness, but we cannot and will not discover a standard of what constitutes something’s being better if we do not analyze what I have called betterness. In addition to the aforementioned concepts which comprise goodness, therefore, the concept of comparison will become central to our discussion.

Before we move on, let us remember our other tasks with which we will have to deal in the next chapter. Recall the Single-Person Problem that I presented in Chap. 3. In presenting this problem, I demonstrated that, in a single-person world, all utilitarian, deontological, contractarian theories, as well as many others, break down because they all suffer from what Slote has called “self-other asymmetry.” I characterized this self-other asymmetry by its dependence on certain logical or other relationships between an individual and his society. I stated, thus, that the fault with these theories lies in the dependence of one’s acting morally (or even rightly or justifiably) on the existence of other beings. On these theories, therefore, if you were the only person, or even the only sentient being, existing in the world, you would have no reason for acting in any particular way. My analysis of goodness will demonstrate that this problem is soluble only through the paradigm of virtue ethics, which stresses *acting well* and not merely *acting morally*.

Recall, also, that we will have to show how the intrinsic goodness we discover from the analysis satisfies the criteria for intrinsic goodness that Magnell has set forth and which I elucidated in Sect. 5.1 of this chapter. As the reader might recall, he states that, in order for “good” to have a semantically predicative role,

it will need to have meaningful application to “a predicate with an extension that necessarily includes as a subclass of its extension the extension of the predicate of the given sentence.”²³ Intrinsic goodness, I will show, does, indeed, fulfill this criterion.

My ultimate goal from all of this analysis is to demonstrate that intrinsic goodness exists but that intrinsic goodness, itself, provides only a vacuous basis for judging what is better and best. Goodness is, indeed, a peculiar property, but not peculiar in the way non-descriptivists and some descriptivists have characterized it in the past. In order for goodness to obtain an intelligible standard of comparison, as many other properties have, we will have to take a slightly different approach. Certainly, we will have to analyze goodness, but, if we are intent on finding a standard of comparison, we will need to analyze *betterness*. This betterness, along with goodness, is a property as well, and we shall see why. Betterness, also, is not merely extrinsic; it is not *based* simply on comparison; it is, in fact, *a type of comparativeness*, although it is still correct to state that *statements concerning what is better* pick out extrinsic goodness. Betterness, I will demonstrate, is intrinsic, but this intrinsicality is not due to goodness alone. To aid the reader, we can, for now, think of betterness as “intrinsic goodness comparativeness” and, thus, intrinsically goodness-comparing. As I will explain, we need to discover not one but two intrinsic properties: goodness and betterness. Both, we shall see, possess a special type of intrinsicality, what I shall call *reflexive intrinsicality*.

²³ Magnell, “Evaluations as Assessments, Part II.”

Chapter 6

Beyond Dialectical Necessity: Assertoric Necessity and the Grammar of Goodness

In this chapter, I demonstrate that, while incorporating important elements of dialectical theory, value theory is indeed ultimately post-dialectical and virtue-theoretical. These dialectical approaches are still sound and widely applicable, I claim, but, as I have been stating elsewhere, it is only a more comprehensive aretaic theory of value which possesses the capacity to fill many of the theoretical gaps which loom large in these models. I demonstrate, via an analysis of the concept of value, informed by all of our previous analyses, that it is a particular sort of virtue theory which answers the questions I have posed throughout this work which I have claimed that non-aretaic models, in general, fail to answer. I achieve this result, first, by demonstrating that value is what I term “reflexively intrinsic,” and, by consequence, that “goodness” can be accurately and adequately defined using the notion of tendential necessity of functions and, subsequently, that betterness can likewise be defined using the notion of comparative tendential necessity of functions. Through the phenomenon of reflexive intrinsicity, I claim, the concepts of value, goodness, betterness, and bestness are inherently teleological and, by consequence, teleologically comparative.¹ I utilize this idea of teleological comparativeness to demonstrate that goodness and value have ultimate foundations in rationality and the Rational Framework, as I have characterized them. As we shall see, rationality constitutes a transcendental conceptual framework which grounds the concept of value itself. It is from the logic of this grounding, I claim, that we can derive all-encompassing answers to even the hardest questions in value theory. The end result is an over-arching, assertoric, aretaic, value-theoretical model which stands grounded on its own but which also acknowledges the usefulness and applicability of the Gewirthian and Habermasian models. This over-arching model, we shall see, is capable of providing answers to some of the profoundest and most personal questions of value.

¹ Special thanks to Thomas Magnell for giving me the idea to more closely analyze and consider the concept of comparativeness with relation to goodness, especially extrinsic goodness.

6.1 Reflexive Intrinsicity and the Teleologically Comparative Tential Necessity of Functions

In this section, I demonstrate that goodness and betterness are intrinsic properties of the world. Their intrinsicity has been difficult to ascertain, I argue, because they are what I call reflexively intrinsic. I also define these concepts in terms of three central concepts: tendency, necessity, and function. In order to understand betterness and bestness, we need to focus on the concept of comparativeness as well and, more specifically, teleological comparativeness. Once we closely examine these concepts, we find an inextricable link to the pseudo-Carnapian concept I have put forth called “the Rational Framework” via the concept of the teleologically comparative tential necessity of functions. It is this Rational Framework, I will argue, that ultimately renders value judgment meaningful and assertorically, inescapably justifiable.

It seems as though I have been introducing so much at once: empirical-linguistic frameworks, reflexive intrinsicity, a seeming ambivalence toward the good-for thesis, and this perhaps dubious-seeming concept of betterness, which, nevertheless, value theorists have historically upheld. All of this, you might be bristling, seems very confused. Let me see if I cannot change your mind. The conceptual links between all of these are intricate and subtle, so we must undertake a careful and rigorous analysis. In this chapter, I will offer a naturalistic account of value via an assessment of the relationships between these concepts. I will demonstrate that, once we examine these relationships, it becomes apparent that certain truths concerning the *summum bonum* derive from statements concerning goodness and value.

6.1.1 Assaying Some Peculiarities

Let us start with betterness. This so-called betterness, one might object, is obfuscatory; why not simply say “extrinsic goodness?” I have chosen to speak of betterness rather than extrinsic goodness for some very specific reasons. The most salient reason concerns the subtle linguistic difference between “good” and most other adjectives; as we will find, it is actually clearer and, thus, more useful to use the term “betterness” instead of extrinsic goodness because, although betterness is indeed extrinsic goodness, it will eventually become clear that betterness is not *merely* an extrinsic property; it is also an intrinsic property. While this duality may currently seem strange, it will be prudent to keep this duality in mind as we progress in this analysis. Such a duality, we shall see, is possible and is uniquely peculiar to intrinsic goodness comparativeness, including also what one might call bestness. To see the consequences of this parlance, let us consider, first, the general nature of most attributively used adjectives which signify extrinsic properties. When we say that something is wider than something else, we are, of course, making implicit reference to the intrinsic property of width. We say, thus, that something is wider than something

else because it has more intrinsic width than some other object. Goodness, however, has a genuine peculiarity in how it functions. Notice that, when someone says that a desk is wider than something else, we have no problem in asking “Wider than what?” and receiving the answer “Wider than that desk over there.” Most of us would be satisfied with this answer, but a few of us might inquire further: “How are you judging that this desk is wider than that one?” “Because,” the person might respond, “this desk has more (intrinsic) width than that one.” Most of us go no further in asking for a definition of width, but we are confident that, if asked for a definition, someone, if not this person himself, would be able to give us a straightforward definition in terms of a set of spatial relationships. This, we should note, is not exactly how “good” functions; it is somewhat different. Let us pose a relevantly similar set of questions using “good.” Someone, first, asserts that a desk is better. Same as with “wide,” we ask “Better than what?” We, then, receive the answer “Better than that desk over there.” Most of us, I think, would not be satisfied with this answer. “Why is it better?,” we would ask. “How are you judging that this desk is better than that one?” Still, many of us probably would not be entirely satisfied with the next response: “Because this desk has more (intrinsic) goodness than that one.” The reason why this line of questioning and response seems so befuddling is not because this is a series of open questions that can never be answered due to some mysterious missing meaning. The reason for this, rather, is that we are asking the wrong questions entirely; we are assuming that “good” works in *exactly* the same way that all other attributive uses of adjectives work. “Good” possesses the peculiarity of having the ability to be used in response to these sorts of aforementioned assertions in a wholly different, yet perfectly meaningful and intelligible, way. We can ask, for instance, “Good for what?” or “Better for what?,” whereas “Wider for what?” and “More massive for what?” usually do not make sense in such a context. Yet we do not judge from this that “wide” is an undefinable word containing some missing meaning. This manner of contextual usage is chiefly why “good” seems more elusive in its meaning; it does not have a relevantly similar function to other sorts of adjectives.

6.1.2 *The Good-for Thesis*

Why, however, does “good” have this function, while other adjectives usually do not? Why do we seem to be able to ask “Good for what?” *ad infinitum*? As mentioned before, Wittgenstein, as well as Thompson, have also asked this question and thus have noticed this peculiarity but seem to have fallen short of providing an explanation for why “good” seems to possess such a peculiarity. Zimmerman, by contrast, in opposition both to Geach and Thompson, has argued, rightly I believe, that it does not follow from the assumption that all goodness is goodness-for that what he calls “generic goodness” does not exist. As he states:

Thomson appears to believe that her First Thesis (that all goodness is goodness in a way) implies her Second Thesis (that there is no such thing as generic goodness). Does it? Consider the analogy with shape. (Color would provide another analogy.) It’s plausible to

say that all shape is shape in a way, that nothing can have ‘pure, unadulterated’ shape. What does this mean? It seems to mean (at least) that shape is a determinable property; nothing can have a shape without having a particular shape. Yet a determinable property is, of course, a property – a property, indeed, which unites otherwise dissimilar things. (Shape is something common to both squares and circles.) Why not take generic goodness to be on a par with shape in this regard?²

He thus proposes an intriguing alternative. If all goodness is goodness-for, and intrinsic goodness exists, then intrinsic goodness, he supposes, might imply “just one way of being good.”³ He follows this, rightly, by asking what could possibly constitute such a way. His answer, quite interestingly, is “*ethical goodness*.” He continues: “When it’s said that beauty, or knowledge, or pleasure, or virtue is *intrinsically* good and that, for example, activities that promote such states are *extrinsically* good, what’s meant is that all these things are *ethically* good.” While interesting and potentially very important, this thesis continues to beg the question, for what, precisely is ethical goodness? It seems that all Zimmerman has achieved is the replacement of the term “intrinsic” with the term “ethical.” Indeed, I believe he overlooks one of Thompson’s principal points, namely that “it is not just happenstance that the word ‘good’ appears in all of those expressions [which she presents]: its meaning surely does contribute to their meanings, and we need an answer to the question how it does.”⁴ Thus, while Zimmerman is close to a resolution, I believe his analysis falls short. In order to explain precisely why goodness seems to function in such a manner, we will have to ask what, exactly, ethical goodness is good for and why, in turn, that which it is good for is good.

6.1.3 *The Reflexive Intrinsicity of Goodness*

The explanation that I shall set forth concerning why all statements concerning goodness are, indeed, statements concerning what is goodness-for is that goodness is reflexively intrinsic. This is due to the central role that what I will call “tendential necessity” plays in “good-for”-statements. No need to bother ourselves just yet concerning what this means; suffice it to say, for now, that tendential necessity is a kind of necessity expressed in judgments of goodness. It is due to this sort of necessity that is inherent in “good” that, just as we can say “X is necessary for Y” or “X has a tendency to yield Y,” we can say “X is good for Y” and, indeed, that we must always, in some sense, say “X is good for Y” when we say “X is good.” Thus, Wittgenstein and Thompson were right in some sense. They were wrong, however, to state that we can never meaningfully say “X is good” by itself, although we are always implicitly referring to some Y. Wittgenstein and others, however, overlooked

² Zimmerman, “Defending the Concept of Intrinsic Value,” 158.

³ *Ibid.*, 162.

⁴ Thompson, “The Right and the Good,” 152.

that this Y, itself, could be *reflexive*, in other words, that it could be goodness itself. If we carefully scrutinize the concept of goodness and discover, from this scrutiny, the definition of goodness, then we can see that goodness has the peculiar feature, via tendential necessity, of reflexivity. In other words, something can be good for goodness or, in other words, good for being good; the possibility must be retained that there exists an entity or entities which have the capacity to be good, and even the best, for instantiating the concept of goodness itself. It is in this way, I will claim, that value is reflexively intrinsic, namely in the sense that there exist certain concepts that are foundational to all conceptions of goodness and which, due to the perpetually teleological nature of goodness as goodness-for, attain the status of being intrinsically, generically good.

So why use the term “betterness?” Again, let us compare “good” to other uses of adjectives. In the adjective “wide,” the standard for comparison to other wide things is already lucid; it is width. In the adjective “good,” however, the standard is somewhat more obscure. Yes, the standard for comparison is, indeed, goodness, but that does not clarify much. Why is this? This is because goodness, yet again, has as one of its essential features tendential necessity, meaning that, although we have discovered the intrinsic property itself, we have not touched upon its peculiar feature, namely goodness-for, and this is what we are truly after. But if we are to search for goodness-for, we will be stuck in an infinite regress. As I will demonstrate, we know what goodness is and what its definition is, but, as we shall see, simply knowing its definition does not help us much. If we were given the definition of width or mass, we would easily be able to pick out which things are wider or more massive, but even with the partial definition “tendential necessity,” we do not know exactly what to pick out for comparison, since such concepts as necessity and tendency inherently have various relationships to many diverse concepts; there is necessity *for* and tendency *to*, respectively; there is never simple necessity or tendency. This seems to create a mess. Which objects have necessity? All objects do. Which objects have tendency? All objects do. All things are necessary for something, even if it is themselves, and all things have tendencies *to* something. We can see, therefore, that it gets us nowhere to ask “Which objects are more necessary than other objects?” and “Which objects have more tendencies than other objects?” because we are always stuck right back in the same position: “For what and to what?” Nothing is wrong with goodness; it is simply that we have been asking the entirely wrong questions for a long time. Let us, then, start asking the right ones. We must start, I believe, by asking the seemingly naïve questions, “What is a sound standard of comparison?” and “What thing or things make goodness intelligibly comparative?” In other words, goodness is peculiar such that we need, first, to find the standard of comparison which yields that which is better or best before we can begin to compare different good things intelligibly. This standard of comparison, we shall see, is *betterness*. It is literally the standard which makes all comparison within the Goodness Framework possible. I hope, then, that the reader has begun to see why this concept of betterness is necessary. It is necessary, I have said, precisely because, unlike most properties, goodness alone renders no non-infinitely-regressive standards of comparison due to its unique feature of tendential necessity. We have, then, to discover such a standard via an analysis of the reflexive intrinsicity of goodness and betterness.

6.1.4 *Tendency, Necessity, and Function*

So where do we go from here? Let us start simple and begin with the basics, namely the definition of goodness. Like many definitions of predicative adjectives, goodness is going to have a very general definition. Consider width, for instance. Its own definition, in terms of a set of spatial relationships, is very general. This is also the case with mass. Apply both of these properties to specific objects, however, and we get more and more specific descriptions concerning these properties in relation to the objects which they comprise. We might say that an object can have intrinsic mass, but we can also be more specific. We can, for instance, speak of a massive planet. This planet, having a specific mass, will interact in a specifically different way with its environment than if it did not have this specific mass. Thus, depending upon what mass it has, and in which environments, its mass will have a very different role. We should expect the same from goodness. When it is applied to very different circumstances, the goodness of a specific thing will play very different roles, whether it is subordinated to another good, whether it is a moral good, whether it is a prudential good, whether it is a signatory good, etc. Like the various different massive objects, however, these different good things share the intrinsic property of goodness. “Good,” then, like some other adjectives, is *pros hen* equivocal.

We have already considered some of the basic concepts that are related to goodness. We stated in Chap. 4 that positive judgments of goodness concerning one’s choices are necessary and sufficient for “must”-judgments concerning one’s choices. What is true of any “must”-judgment concerning one’s choices, therefore, is also true of any positive judgment of goodness concerning one’s choices. The reader might recall that we demonstrated this by analyzing the structure of “ought”-judgments, finding them all to be regulated by an underlying, ultimate, “must”-judgment. One might object that one might state that one’s choice is the better one and, thus, imply that one ought to make that choice but, however, that it does not follow that one *must* make that choice. I anticipated this objection arising again, and I have already answered it in Chap. 4. I encourage the objector to see this chapter, as well as Figs. 4.1, 4.2, 4.3 and 4.4. “Must”-judgments concerning one’s choices, indeed, are necessary and sufficient for positive judgments of goodness concerning one’s choices, even if this positive judgment of goodness is something as simple as “My action X is good for Y.” The point, here, is to point out, as many others have, the concept of necessity that is inextricably bound up in the concept of goodness. Foot, Anscombe, and Aristotle himself have acknowledged, if at least implicitly, the relevance of necessity to goodness.⁵ Let us consider another possible bi-conditional relationship; let us see if we can leave choice out of it altogether and say something more general. What do “must”- or “ought”-judgments not concerning choice, alone, entail? What do we imply, for instance, when we say such things as “The sun ought to or must be green instead of orange,” “It ought to or

⁵ Nancy Sherman, *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. (Cambridge: Cambridge University Press, 1997).

must rain,” “The house ought to or must be bigger,” or “I ought to or must eat?” All such judgments, in fact, *are* value judgments; they *are*, covertly, judgments of goodness. They do not, for instance, simply *entail* judgments of value or goodness. Piggybacking on Smith’s thesis that value judgments are statements of belief, we can see how we could roughly translate such statements as “The house ought to or must be bigger” and “I ought to or must eat” to statements which instantiate goodness. For instance, “The house ought to or must be bigger” can be roughly translated to “The house’s being bigger is a necessary condition of fulfilling goodness-standard/value X,” and “I ought to or must eat” can be roughly translated to “My eating is a necessary condition of fulfilling goodness-standard/value X.” Let us turn to some other judgments of goodness. Consider the following statements: “It would be better if the sun were green instead of orange,” “It would be better if it rained,” “It would be better if the house were bigger,” and “It would be better if I ate.” Then, try to say that it is not the case that these hypothetical states of affairs should not or must not be. This, however, does not make a lot of sense, unless, of course, we interpret this in the following way. Certainly, we will sometimes hear people say such things as “It would be better if the house were bigger, but it should not be bigger,” but we should be wary of the meaning of such statements. One possibility is that this sort of a statement represents a kind of backsliding. The statement could thus be something more like “It would be better if the house were bigger, but then again it would not be better” or “The house ought to be bigger, but then again it ought not to be bigger.” As my analysis in Chap. 4 has shown, “ought”-judgments have correlative “can”- or “cannot”-criteria that they have to fulfill if they are to proceed to “must.” If they do not fulfill these criteria, they are not judgments of categorical obligation. When someone says “It would be better if the house were bigger, but it ought not be bigger,” someone might, alternatively, be implying “It would be better if the house were bigger if or only if condition X were met, but since condition X has not been met, it ought not to be bigger,” where condition X is something along the lines of “if I did not have to spend money on some other very important affair.” Thus, here, what is stopping the “ought”-judgment from being a judgment of categorical obligation is the “can” of permissibility. It could, of course, alternatively be a “can” of possibility. Keep in mind, however, that these sorts of judgments constitute indirect references to categorical obligation. In the above example, the reason why the person judges that it would be better if the house were bigger but that it ought not to be bigger is that a “can” of permissibility has not been met. And what is a can of permissibility? It is a “must.” When we say that someone’s judgment has not passed his own “can” of permissibility, we are implying that he has made an underlying “must”-judgment or, truly, a “must-not”-judgment that prevents his “ought”-judgment from being categorically obligatory. Thus, we can see why such judgments as “It would be better if the house were bigger, but it must not be” make no sense, that is, unless we interpret the “must,” here, as a weakened “must.” In other words, if we judge some state of affairs to be better or best, it is tantamount to judging that such a state of affairs must be *if it were possible that it could be or if it were permissible for it to be*. As I stated in Chap. 4, there is only obligation and non-obligation; there is nothing in

between. It seems like there are judgments in between these two poles only because of the manner in which “ought”-judgments mediate “must”-judgments. Thus, at the heart of judgments of comparative goodness concerning states of affairs is a direct or indirect judgment of categorical obligation and vice versa. This, however, is the catch: obligation for *what*? And what makes this obligation *categorical*? We seem to reach, here, a similar aporia as with goodness and, as we shall see, there is a very apparent reason why. In the meantime, let us at least remark that such a question problematizes deontology, for this analysis suggests evidence for the claim that deontologists may not merely assume a conception of justice or duty without answering certain seeming questions of value, namely “Justice for what?” and “Duty for what?” Thus, for instance, Habermas and Forst may not eschew the good in favor of the right without first appealing to the good.

Before moving on to an analysis of the concept of categorical obligation, it is worth mentioning that this analysis also stands the test of Foot’s classic criticism of the claim to an essential connection between goodness and choice.⁶ To be clear, Foot’s arguments with regards to this issue were never directed against the sort of argument I am proposing. Instead, she rightly lodged her arguments against those who claimed that all statements of goodness necessarily imply statements of choice and vice versa. Certainly, I am not claiming that expressions such as “good eyes,” “good roots,” or “good teacher” imply an assertion of choice, that is, so long as these expressions are being used non-intrinsically and have no implicit logical connection to choice. Allow me to briefly explain this. Foot’s own claim had roughly been that such expressions imply no imperative to choose these things. This is certainly true, but the two conditions I have mentioned are vital to making this claim more precise. While Foot rightly criticizes philosophers who make the claim that statements of only evaluative uses of “good” imply an imperative to a certain choice concerning the good entity in question by, in turn, calling into question the evaluative-non-evaluative distinction with regards to goodness, she regrettably overlooks the idea of value intrinsicity. This idea is an important one, for, while it is the case that such expressions as “good eyes,” “good roots,” and “good teacher” are evaluative, these expressions may also be used to denote intrinsicity. If it is the case, therefore, that I am judging good eyes to be intrinsically good in that they are tied to the fulfillment of some *summum bonum*, ultimate principle, or otherwise intrinsically good *telos*, then there are certain manners in which these expressions would necessarily imply an imperative to choose the states of affairs tied to these expressions. If, for instance, I state “Having good eyes is better than having bad eyes,” “A tree’s having good roots is better than a tree’s having bad roots,” or “It is better to have a good teacher rather than a bad teacher teaching our son, and learning is intrinsically good,” and these statements are true, then, recalling that we are uttering these statements with reference to intrinsic goodness, so long as no event or occurrence of greater intrinsic goodness

⁶ Philippa Foot, “Goodness and Choice,” *Virtues and Vices: And Other Essays in Moral Philosophy*. Ed. Philippa Foot. (Oxford: Oxford University Press, 2002): 132–145.

outweighs these goods in a particular instance, it is logically absurd to state that it is not the case that we must promote the possession of good eyes, trees' development of good roots, and our son's access to a high-quality teacher. Foot's argument, by contrast, merely addresses the expressions "good eyes," "good roots," and "good teacher" as they are used with reference to non-intrinsic *teloi*. Good eyes, good roots, and good teachers might, by definition, be eyes that do well that which eyes do, and so on for the others, but it does not follow from this that the good functioning of eyes is good. If one proceeds to show that the good functioning of eyes is, indeed, good, then one can justifiably state that there is an implicit link to the categorical obligatoriness of one's choice with regards to the promotion of good eyes. This, however, will be a rather superficial categorical obligatoriness, for one must further prove that, given a conflict between good eyes and, say, good roots, good eyes are *better*. This kind of statement provides a far clearer and more direct link to the categorical obligatoriness of the choice to promote the development or prevent the deterioration of good eyes. Now that we have cleared this potential red herring, we may move on to a further analysis of the connection between goodness and obligation.

Obligation, we have seen, is essentially necessity. If some entity Q is obligatory for some entity P, then it is necessary for that thing. Likewise, if some entity Q is necessary for some entity P, then it is obligatory for entity P or, to put it another way, it is obligatory for entity P to have entity Q *qua* entity P itself; it might, for instance, not be *categorically* obligatory for entity P to have entity Q, but this is a different matter altogether. Thus, we can see that, perhaps, a better word to use in the aforementioned examples, instead of "obligation," is "necessity." If I state "It would be better if the house were bigger," I am concurrently making some judgment of necessity such as "It would be necessary that the house be bigger if or only if some 'can'-criterion were met;" if the reader is confused by how a congruence is possible between a judgment of necessity and an "only-if"-clause, I encourage him to see my comments on this in Chap. 4. Yes, I have made some judgment of necessity, but what makes it *categorically necessary*? I stated, in Chap. 4, that the person *judges* something to be categorically necessary because of some standard of goodness that she has. What this means is that categorical necessity is *necessity for* such a standard or the upholding of such a standard. Such a judgment, however, is not actually categorically obligatory if the standard is unsound. But how do we prove the standard to be sound? This will come eventually. For now, suffice it to say that the concept of necessity, as many other theorists have either shown or implied, is central to the concept of goodness.

What else is central to the concept of goodness? We know that function is central to it. This has been clear not only from Foot's analysis, but from the analyses of other virtue theorists such as Geach, Slote, Aristotle, Epicurus, and numerous others. It seems obvious enough that, when we refer to the goodness of an entity, we are referring, in some broad sense, to its function or to the function which constitutes its essence. A good X, thus, is an X which does well that which Xs do; it is an X which functions well as an X. Thus, in every judgment of goodness, we are citing both function and necessity. How, exactly, are we citing them though? To

begin to discover this, let us return to the good-for thesis. If something is good or better, it is always good for something or, in the case of “better,” better than something else for something. We might say, in other words, that, if something is good or better, then its function is good for some other function or that its function is better than some other function for some function. It is here that we must be careful. We do not, for instance, want to state that “Function X is good for function Y” means “Function X is necessary for function Y,” and we certainly do not want to state that “Function X is better for function Y than function Z” means “Function X is more necessary for function Y than function Z.” Why do we not want to state this? Because it is not true. Yes, the concept of necessity is central to the concept of goodness in some way, and, yes, we always cite *some sort* of necessity when we make judgments of goodness, but we do not always cite necessity *in this specific way*. For instance, it is sensible to state, “If function X is necessary for function Y, then function X is good for function Y,” but it seems nonsensical to state “If function X is good for function Y, then function X is necessary for function Y.” What is going on here? Why do these expressions lack bi-conditionality? Since it is clearly not the case that every judgment of goodness entails a judgment of the kind of necessity in the aforementioned examples, we might be tempted to conclude that judgments of goodness entail a judgment of sufficiency, but this, too, would be wrong, as I will explain. What, then, do statements of goodness entail if, as I claim, all statements of goodness cite some sort of necessity? In order to find the answer to this, we have to make a temporary departure from the realm of deductive logic and step into that of inductive logic.

In various past analyses, it has been suggested that something that is good for something else is conducive to that thing and vice versa. Conduciveness, however, is a vague and flimsy concept. It is, however, somewhat useful. What does it mean for something to be conducive to something else? We might be tempted to state that something is conducive to something if it is sufficient for it. To prove this, we might use expressions associated with conduciveness such as “to lead to” or “to yield.” The latter part of this makes a lot of sense; to be conducive to, under one interpretation, means “to yield” or “to lead to,” and these expressions seem highly relevant to the concept of goodness. But these, by themselves, are inadequate. First, let us see the flaws inherent in assuming that these expressions imply sufficiency. If some function X is sufficient for some function Y, one might state, then it is conducive to that function. One might state that sufficiency, like conduciveness, expresses X’s “leading to” or “yielding” function Y. One might also misapply Kantian epistemology and cite sufficiency and necessity as representative of a cause-and-effect relationship. Strictly speaking, however, to speak of sufficiency as “leading to” or “yielding” some function is inaccurate. Some function can be sufficient for some other function without being conducive to it or “leading to” or “yielding” it or anything else for that matter. Likewise, some function can be conducive to some other function without being sufficient for it. Let us take the example of existence and life. Existence, we would say, is necessary but not sufficient for life. Thus, we would say, life is sufficient but not necessary for existence; we are speaking of existence, in general, of things that are not necessarily

living such as gases and planets. It makes sense to say that existence, being necessary for life, is conducive to life, even though it does not *necessarily* lead to or yield life, and this will be explained momentarily. By contrast, it does not make sense to state that life, being sufficient for existence, leads to or yields existence. Life is not even remotely necessary for existence; minerals, for instance, are not life, yet they exist. What, then, does “conductive to” mean? Conduciveness, we can actually see, is an inductive concept pertaining to probability and, ultimately, tendency. If we import this inductive concept of tendency into deductive logic, we actually find ourselves doing deontic logic; this is important if we consider how closely interrelated are deontic modal operators such as “must” and value concepts. If we think of “conductive to,” and, thus, “leading to,” “yielding,” and, ultimately, “good for” as deontic concepts, we can clarify the meaning of these concepts. We shall see momentarily why they can be interpreted as deontic concepts. If something is conducive to some function, it does not follow that it leads to or yields that function. It follows, rather, that it has some tendency to lead to or yield that function. In fact, we can make a deductive move here. If some function X that is conducive to some function Y tends to lead to or yield Y, then it is *necessary* for some function X’s being conducive to some function Y that function X not be non-conductive to Y; or, in other words, it is *necessary* that it not be impossible for function X to lead to or yield function Y. If we translate the “necessary for” here to a “must,” we see that this is a clear deontic relationship.

If we take the example of existence and life, once again, we find that existence is conducive to life because it has some tendency to lead to or yield life because, first and foremost, it is not impossible for existence to lead to or yield life. This, we see, is much more substantive and explanatorily powerful than simply stating that existence leads to or yields life since, strictly speaking, it does not; it simply has some tendency to lead to or yield life. Compare this with life’s supposed conduciveness to existence; life does not, strictly speaking, have a tendency to lead to or yield existence. Life does not, in other words, bring about existence. It is impossible for life to bring about existence; rather, it is only possible for existence to yield life with its already-existing particles which comprise organic molecules, which comprise cells, and so forth. To say that such things have a tendency to yield existence is nonsensical since this is tantamount to claiming that cells, for instance, have a tendency to yield existence or, alternatively, that cells have a tendency to yield themselves.

We might, however, state that my hand’s moving a book is sufficient for the book’s moving but not necessary for its moving, and, thus, we might say that the function of my hand’s moving in a certain way is good for the function of the book’s moving. Although this seems like a rather awkward use, and perhaps an abuse, of the word “good,” let us, for the sake of argument, run with this. If we accept such an example, it would seem that there are certain instances in which sufficiency can be used in “good-for”-statements. This thinking, however, is flawed. Let us revisit the concept of conduciveness. We said that it is a necessary condition of conduciveness that function X not be non-conductive to function Y. We can conclude, therefore, that, even in such sufficiency examples, there is necessity implied here, namely that

it is necessary that function X not not-have a tendency to yield Y. In other words, if we want to say “My hand’s moving in a certain way is good for the book’s moving,” we must admit that it is necessary for my hand’s moving the book (not the book’s moving) that my hand have some tendency to move the book, since “My hand’s moving in a certain way is good for the book’s moving” entails “My hand’s moving in a certain way necessarily has some tendency to effect the book’s moving.” Even in such statements which hypothetically entail sufficiency, there is still a necessary condition concerning the tendency of X to yield Y. Necessity, then, indeed *is* central to statements of goodness, but it is a specific sense of necessity that is central: that of the necessity of tendency. This necessity of tendency is what I shall call *tendential necessity*. Thus, in order for function X to be conducive to and, thus, good for function Y, then it is necessary that X have a tendency to lead to or yield Y, and this is important.

Gewirth has made the mistake of speaking of something’s being “more necessary” than something else, but this is, strictly speaking, inaccurate. A function is either necessary for or not-necessary for; there are no “degrees of necessity.” Gewirth, however, is on the right track. Instead of “degrees of necessity,” he should speak of “degrees of greater or lesser tendency to lead to or yield.” Necessity is still central here, namely tendential necessity, but it would be fallacious to measure some function X as better for some function Y than some function Z is for Y because X is, somehow, “more necessary” for Y than Z is. It would be more accurate to state that some function X is better for Y than Z is because X has a greater tendency to lead to or yield Y than Z does. We must, therefore, be wary of making this kind of mistake.

6.1.5 *Applications to Goodness*

We must be careful which functions we are ascribing tendency to which others, for it is not always clear in natural languages how we are to interpret an expression. For instance, we can give the classic example of “good for a thief.” This expression could entail one of three things. It could imply either (1) those properties which we ascribe to a thief which make him function to a prescribed *telos*; or (2) those properties which make the thief capable of fulfilling his interests, however we may describe these; or (3) those properties which make the thief intrinsically or morally good. Although most people would not usually refer to (3), we cannot rule it out as a possible implication of “good for a thief.” The distinction between these different implications is one of class. In uttering the criteria in (1), one implicitly refers to the class of things related to the essence of thief-ness. What is the *telos* of thief-ness? This *telos*, of course, depends on the essence of thief-ness, which we would have to carefully ascertain, since there are many kinds of thieves. We could, tentatively, say that the essence of thief-ness is “attempting to steal” and that, therefore, the *telos* of thief-ness would be “stealing successfully,” where “successfully” merely means “with as little prevention of or intervention in the stealing as possible.” In order to find out what is

good for a thief, in this sense, we would ascertain the function, and therefore the essence and *telos*, of thief-ness. Then, we would attempt to discover those things which are necessary for the tendency to yield this function. To ascertain what is *better* for a thief, in this sense, it would be necessary to discover those functions which have a greater tendency than other functions to yield, first, the essence and, then, the *telos* of thief-ness. The thief which has his *telos* fulfilled is a good thief according to (1). In uttering the criteria in (2), we are referring to the class of things which are in the interests of a thief, specifically. "Interest" is a very malleable word. Although it is typically distinguished from desire, interest is also often confused with desire. Interest, as I and many other theorists conceive it, is a moral term which expresses the good of the person in question, a good that is independent of that person's desires. On this definition, presumably, it would not actually be in anyone's interests to be a thief in the first place, *ceteris paribus*. On the desiderative conception, however, which is typically spoken of as self-interest, we would simply take into account the thief's desires in order to determine his interests. On this account, it is in the thief's interests not to die, not to be maimed or injured, not to be caught or tried, and, of course, to steal as much as he can of what he desires. To determine what is good for a thief, in this sense, we would have to, again, determine the essence of his thief-ness, which is stealing and his *telos* as a thief, which is stealing successfully. However, this time, the emphasis is specifically on the interests of the thief. Since, as we have acknowledged, these interests are self-interested desires, all we would have to do is discover what is good for the fulfillment of these desires, which is the *telos* of the desires. In other words, we would have to discover that which is tendentially necessary for the fulfillment of these desires. As in (1), we must first discover what is necessary for the interests, themselves, since, again, nothing can have a tendency to yield a given function if it is not a necessary condition that there be such a tendency to yield that function. A necessary condition of there being a tendency to yield those interests, we might say, are the thief's desires. Let us be clear that we are merely saying that the desires are necessary for themselves, since the desires and the interests here are the same. The desires, additionally, are necessary for the tendency to yield the thief's *telos*, the fulfillment of those desires. Thus, since the desires are necessary for these interests, the desires are good for these interests, and, because the desires have a tendency to yield the fulfillment of these interests, the desires are good for the fulfillment of these desires; they are, in fact, necessary for such fulfillment of desires. Desiring, while necessary, however, is not the only function that is necessary for the fulfillment of the desires. Thus, while the desires are good for this end, desiring is not the only function which is good for this end. Some other necessary conditions of being good for this end of fulfillment of the thief's desires include, as highlighted above, not dying, not being maimed or injured, and not being caught or tried. Prudence, therefore, among other functions such as vigilance and self-control, are necessary or at least tendentially necessary conditions for the *telos* of fulfillment of the thief's desires. In other words, if the thief is to fulfill his desires, he will need to be prudent, vigilant, and self-controlled. This is simply a rough picture of what the conditional statement would really be; in other words, there might be further necessary conditions for this *telos*. In any case, the example is instructive. What we are

beginning to see is that these functions, namely prudence, vigilance, self-control, and others, are special types of goods. They are, in fact, virtues. Virtues, we can see, are specific types of functions, which we may also call character traits or character states, which are necessary, sufficient, or sometimes merely tendentially necessary for a specific *telos*. Virtues, it must thus be acknowledged, have a relationship to their *telos* characterized by necessity but, more specifically, the necessity of specific tendencies to the *telos*.

It is now that we can move on to interpretation (3). When a person utters the criteria in (3), she actually implicitly refers to the superclass “person,” although she explicitly refers to the subclass “thief.” On this interpretation, we must judge what is good for a thief based upon, as I previously stated, “those properties which make the thief intrinsically or morally good.” Most of us would acknowledge, however, that a thief is not morally good because “thief” does not belong to the extension “things that are morally good.” Its superclass, “person,” however, does belong to such an extension. In making this specific judgment concerning what is a good thief, then, we are really making a judgment about the person and what makes the person good. Just as with a thief, we would have to find the essence of a person, which, as we already know, is rationality. We would, then, have to find the *telos* that follows from this essence, which, as I will prematurely state, is to do what is better or best. How such a *telos* follows from the essence of rationality we will see. Needless to say, the *telos* of doing whatever one wants as long as it is done through rationality does not follow from the essence of rationality, which I shall show.⁷ Furthermore, I do not believe a person a person to be, as many eudaimonists seem to claim, a bad or vicious person if she has, for one reason or another, failed to achieve *eudaimonia*. As Robert Heinaman and others explain, this is a misinterpretation of Aristotle and is, furthermore, untenable.⁸ This position also grossly overlooks other classic *teloi*, such as *ataraxia* and even *nirvana*, and dismisses them from possible consideration. As my analysis progresses, we shall see exactly why all of these are dangerous oversights and pitfalls.

6.1.6 A Naturalistic Definition of Goodness

Using the three examples I have given above concerning the different connotations of the expression “good for a thief,” we have begun to make some discoveries. First, we have seen that, in all three examples, and thus, in whichever way we want to describe something as good for something else, the aforementioned concepts of function, necessity, and tendency are elegantly applicable and, yes, interchangeable

⁷ Alfonso Gómez-Lobo, “The Ergon Inference,” *Phronesis* 34, no. 2, (1989): 170–184.

⁸ Robert Heinaman, “Rationality, Eudaimonia and Kakodaimonia in Aristotle,” *Phronesis*. 38, no. 1, (1993): 31–56; Timothy D. Roche, “In Defense of an Alternative View of the Foundation of Aristotle’s Moral Theory,” *Phronesis*. 37, no. 1, (1992): 46–84.

with the expression “good for.” But what, exactly, is interchangeable? So far, it seems, I have used these concepts only in terms of each other and still slightly disconnected from each other. What expression, then, is interchangeable with “good?” In other words, what is it that “good” means? Let us start with “good for.” We found that, when we use the expression “good for,” we are at once implying necessity, tendency, and function. We have seen, however, that we have been ordering these concepts in a very specific way. We have said, for instance, that, in order for *function X* to be good for *function Y*, it is not the case that *X* must be necessary for *Y* in the sense of *X*’s being a necessary condition of *Y*, but, rather, that *X* must be *tendentially necessary* for *Y*; in other words, it is a necessary condition of *X*’s being good for *Y* that *X necessarily have some tendency to lead to or yield Y*; furthermore, it is a necessary condition of *X*’s being good for *Y* that *X not render impossible the tendency to lead to or yield Y*. Thus, function *X* has two necessary conditions in order to be good for *Y*, both of which entail each other. In addition, we must remember that, since any function *X* that is necessary for *Y* necessarily has *some tendency to lead to or yield Y*, anything *X* that is necessary for *Y* is good for *Y*. We begin, thus, to see the establishment of a definition. If some function *X* is tendentially necessary for some function *Y*, then function *X* is good for function *Y*. Conversely, if some function *X* is good for some function *Y*, then function *X* is tendentially necessary for function *Y*. This can be seen not only from past analysis but also from the following set of assertions. Consider the following assertion: “Green tea is good for the prevention of cancer.” We can, in fact, use any variant of the term “good,” whether it be “valuable,” “important,” “useful,” etc. Consider, then, the following assertion: “It is not the case that green tea is tendentially necessary for the prevention of cancer” or, if you like, “It is not the case that green tea has the tendency to yield the prevention of cancer.” Let us examine the two assertions “Green tea is good for the prevention of cancer” and “It is not the case that green tea has the tendency to yield the prevention of cancer” side by side. Are they consistent? Let us find out. To say that green tea is good for the prevention of cancer is to say that green tea has some property or properties which, which consumed, contribute in some way to the prevention of cancer. In other words, green tea affects or, rather, *does something* to the body which makes it *less probable* or *which lessens the tendency* that that body will develop cancer in the future. What else can we conclude from this but that green tea is, in fact, good for the prevention of cancer? Has the function of green tea not affected the function of cancer in some way such that cancer was less likely to develop? Perhaps it is not the case that green tea is good for the prevention of cancer, but then we will have to admit this using the same criteria: tendential necessity of yielding the prevention of cancer, as well as the empirical relationship of the function of green tea and the function of cancer to each other. What other criteria could we possibly use? It seems, thus, that we have exhausted the criteria we can possibly use to judge something’s goodness for something else and, secondarily, that there is no mysterious missing or hidden meaning in the expressions we are using. If we were to subject this expression to the open-question test, we need only answer “Because of green tea’s tendential necessity for the prevention of cancer” to the supposedly open

question “Why is green tea good for the prevention of cancer?” in order to show that the whole open-question argument is fallacious. If the aforementioned assertions “Green tea is good for the prevention of cancer” and “It is not the case that green tea has the tendency to yield the prevention of cancer” are inconsistent, which I have adequately shown they are, then we see the clear grounds, now, for the following bi-conditional statement:

$$\begin{aligned} & ((\text{TendentiallyNecessaryFor}(x, y) \wedge ((x = y) \vee (x \neq y))) \\ & \leftrightarrow (\text{GoodFor}(x, y) \wedge ((x = y) \vee (x \neq y)))) \end{aligned}$$

where x and y are functions. “Good for,” therefore, literally means “tendentially necessary for.” “Goodness for,” thus, can be defined as “the tendential necessity of functions for,” and “goodness,” itself, can be defined as “the tendential necessity of function.”

Smith has established what he calls a “broad naturalism,” accounting for goodness as a natural property but resisting reductionist approaches.⁹ This account, however, is predicated on the assumption that theorists must undertake analyses of natural properties exclusively via network analysis. He points out the permutation problem inherent in network analyses, demonstrating that certain reductionist accounts yield vacuous descriptions. I believe, however, that my approach avoids the permutation problem for, unlike Smith’s example using colors, goodness is not a particular type of sense experience or qualia. Instead, it is a particular type of relationship founded on the three concepts of tendency, necessity, and function. Furthermore, unlike particular colors, it is the *only* natural property which is founded on these three concepts *qua* the tendential necessity of functions.

6.1.7 First Proof of Intrinsic Goodness

Moving on, how do we know that this is the definition of *intrinsic* goodness? We can begin to discover whether this is the case by asking, aside from grammatical requirements, in what intrinsic goodness would consist. I am in tentative agreement with such theorists as Torbjörn Tännsjö that we must achieve a “concrete” view of intrinsic goodness in terms of treating things in the world as intrinsically good, rather than “facts, propositions, or states of affairs,”¹⁰ but we must understand what I am saying when I state that “the tendential necessity of functions” is “intrinsically good.” The assertion I am making here is of identity and not of predication. I am, therefore, stating “the tendential necessity of functions is intrinsic goodness” and not (not yet at least) “the tendential necessity of functions is intrinsically good.” We have already explored the identity statement “the tendential necessity of functions is goodness,” but it will take a different approach to prove that the tendential necessity of functions

⁹ Smith, *The Moral Problem; Ethics and the A Priori*.

¹⁰ Torbjörn Tännsjö, “A Concrete View of Intrinsic Value,” *The Journal of Value Inquiry*. 33, no. 4, (1999): 531.

is intrinsic goodness. To prove that something—anything—is intrinsically good, one has to prove that that thing is, as you might recall, descriptive-evaluative. That means, on at least a superficial level, that it must be (1) a claim; and (2) predicatively used. It is here that I do not agree with Tännsjö's claim, namely because I do not believe that such things as states of affairs and concepts are not "concrete" and cannot be analyzed for possible intrinsic goodness. Even the intrinsic property of width, for instance, is described in terms of its spatial relations and behavior, which, we can easily see, are both highly conceptual and states of affairs; width is a model, as are such properties as mass and charge. The fact that they are models does not, however, make them abstract; this, however, is a more substantive discussion than can be had here. Intrinsic properties, we see from such other examples as width, mass, and length, are those which necessitate no comparison whatsoever for their definition but are simply defined in terms of their own characteristics and the relationships inherent in themselves; they do not need to be defined in terms of a set of relationships they have to other things in the world which also have the property. And so it is with intrinsic goodness. In order to define it, we need to define it in terms of its own characteristics and its own relationships to itself but, as we have seen, we must do so in a slightly different manner. As we have acknowledged before, good is peculiar in that goodness is always *good for* something; but rather than a hindrance, this peculiarity greatly helps us. What, we should ask, is good for goodness? Perhaps we should be even more audacious: what is best for goodness? We must remember, before we proceed, that anything that is necessary for goodness is, by definition, good for goodness. Would not something that is good for goodness be, quite simply, good? Asked a different way, would not something that is categorically necessary for goodness be *categorically, necessarily good*? The answer, we must admit, is yes. But what such thing is so categorically necessary or good or best for goodness? Why, we must admit, the answer is goodness itself. Think of it this way: "What function is tendentially necessary for the function of goodness?" We will have to answer, if we accept the statement $P \rightarrow P$, that "the tendential necessity of functions is tendentially necessary for the tendential necessity of functions" and, thus, "the tendential necessity of functions is tendentially necessary for goodness," "Goodness is tendentially necessary for goodness" and, finally "Goodness is good for goodness," "Goodness is best for goodness," and "Goodness is categorically necessary for goodness." Where do we get "the best" from? We get it from the consideration that, if goodness is categorically necessary for goodness, and if there is, thus, nothing else that can be good without goodness, then goodness must be the best thing there is for goodness. If, then, the tendential necessity of functions is not intrinsic goodness and is not, furthermore, itself intrinsically good, then it seems that it is not logically possible for anything else to be intrinsic goodness. This peculiarity of which I have long been speaking is what I have called reflexive intrinsicity. While it is certainly the case that all intrinsic properties are defined merely by the relationships they have to themselves and not by their relationships to other things which share the property, it is not the case that they all necessitate proof of their intrinsicity via self-reference, self-entailment, and tautology. The reason for this sort of self-proof of intrinsicity is due to goodness's peculiarity as goodness-for, which is, in turn, due to the

phenomenon of tendential necessity, which is, in turn due to the phenomenon of necessity itself. This entire phenomenon of self-proof is reflexive intrinsicity, and goodness, therefore, as an intrinsic property, is reflexively intrinsic. This phenomenon of reflexive intrinsicity, we shall see, arises elsewhere as well. Although this process is tautological, the process is absolutely instrumental to our understanding the grounds upon which we are justified in claiming that something is good. This process, as we shall see, partially contributes to what some philosophers have posited is a grammar of goodness. This will be explored more in depth later, but for now, it will be important to keep this in mind.

6.1.8 Second Proof of Intrinsic Goodness

Assertions concerning the tendential necessity of functions, as we have shown, are clearly claims and, thus, can be used predicatively. Such kinds of claims can express such value judgments as “Green tea is good/important/useful for the prevention of cancer.” This, we must admit, however, is simply an instrumental use of the term, although instrumental judgments of value are still judgments of value. Now that we have discovered intrinsic goodness, itself, how can we employ it to make other sorts of “good”- and not simply “good-for”-judgments? As I stated before, it does not seem as though there is any clear standard by which we can, as yet, compare different things that are intrinsically good; all we can do is compare the tendential necessities of two or more different functions for some other function. Although we have determined that goodness is good because it is tendentially necessary for itself and, thus, reflexively intrinsic, we have not, in other words, discovered a way of comparing two or more different functions’ tendential necessities for being *better*, which, if one is fooled, one might think is defined as which function is more tendentially necessary for goodness or, if you like, the tendential necessity of functions. Now this is certainly curious. It seems that many properties have a built-in feature that makes it possible, from the start, to engage in comparison with other things that have that same property. Goodness, however, as we have already seen, is peculiar in that it is one of the few properties which, when signified by an adjective, is represented, at least always implicitly, via the expression “good for.” Goodness, it seems, has no such built-in feature of comparison. The only “best”- or “better”-judgments that we can make using the concept of goodness, itself, it seems, are very limited. The only “best”-judgment we have found, we have seen, is that of “Goodness is the best thing there is for goodness because goodness is categorically necessary for goodness.” The tautological nature of the assertion becomes clearer if we re-phrase “best” as “most good.” One “better”-judgment we can draw from goodness itself includes “Goodness is better (more good) than non-goodness.” The other “better”-judgment we can make simply from a derivation from goodness itself is “Existence is better than non-existence,” and this is because we can make the predicative judgment “Existence is good.” Why can we do this? Let us consider this: how would the tendential necessity of functions be possible in

the first place if there were not objects that existed that could have the tendential necessity of functions? Existence, it seems apparent, is necessary for anything to be good; good things, in other words, must have existence to be good. It is also apparent, then, that our previous statement, "Goodness is the best thing there is" is accurate, but not entirely. Yes, goodness is the best for goodness, but existence is also the best for goodness. How does this work? When we state "Goodness is the best for goodness," we simultaneously instantiate the existence of goodness and, thus, necessarily imply that existence is the best for goodness. So when we state "Goodness is the best for goodness," we are actually covertly stating "The existence of goodness is the best for goodness." This is yet another component of the so-called grammar of goodness.

The problem, however, comes when we get out of all of this messy, although very important, abstraction. Yes, existence is good, and it is better than non-existence, and it is the best for goodness and, therefore the best thing there is, but we do not simply want to compare existence and non-existence or goodness and badness in the abstract; we want to compare good things that are in existence and not existence or goodness itself. Let us consider this point. When we state "Existence is the best thing there is," we are not really saying much since existence is, truly, the *only* thing there *is* or, if you like, the only thing that exists. "Existence is the only thing that exists," however, is at best a tautology, and an extremely uninformative one at that. "Existence is the best thing that exists," therefore, is equally as vacuous and uninformative. While a bit tiresome, it is, indeed, momentous, since it demonstrates that something can be intrinsically good, and we should briefly examine this. Let us recall Magnell's requirement that any predicative use of "good" must pick out "a predicate with an extension that necessarily includes as a subclass of its extension the extension of the predicate of the given sentence." Let us take, then, the example I have given above: "Existence is good." This is a true statement, but does it fulfill Magnell's criteria? Yes. Consider what "Existence is good" entails. It entails the expression "good existence." If "good" has an extension that necessarily includes as a subclass of its extension the extension of the predicate "existence," then "good," as used here, is predicative. Let us test this using two methods. In the first method, we will separate the adjective and the predicate and see if they split up logically into a pair of predications and, thus, apply meaningfully to the subject "this." The next method we will use is an analysis of each term's extension.

Let us try the first test. If we separate the terms in the expression "good existence," they do, indeed, split up into a pair of predications. "This is good" and "This is existence" split up and meaningfully apply to the subject "this." How do we know this? Let us proceed to our second test. If we examine each term's extension, we find "good" to be extensionally determinate and intensionally closed. "Good" is a term that has, in its extension, the extension "things which exist," since it is a necessary condition of something's being tendentially necessary for something that something exist which has some tendential necessity; all things, we see, have tendential necessity for an almost innumerable amount of things. Indeed, we can state that all things which exist are tendentially necessary for tendential necessity itself. We can, thus, say "this is good," whatever "this" may happen to be. "Wait a second," the reader might

exclaim, “If everything in existence is good, then isn’t it the case that nothing can be bad or evil? And what about comparing goods? If existence is the best thing that exists, then how is comparison of different goods possible?” These are wonderful questions, and we will take them up with vigor.

6.1.9 The Universality of Goodness and the Comparison of Goods

Let us take the first question. The answer to this question is a qualified yes. Why a qualified yes? It is qualified because all entities that exist are, indeed, good, but here is the qualification: some entities have a lot of goodness, while other entities have an extreme dearth of goodness, although they still have some goodness. The entities that have this extreme dearth of goodness we call bad or, sometimes, evil. We do not refer to them in this manner, however, because they have no goodness whatsoever; they have goodness, but, comparatively, just very little of it. Everything in existence, physical and non-physical, has goodness; everything is good. This might remind us of Leibniz’s “best of all possible worlds” (and Voltaire’s subsequent criticisms), but my views are very different. The key, here, is comparison. This should satisfy the next question. “If existence is the best thing that exists, then how is comparison of different goods possible?” Existence, we are now in a position to clarify, is not, strictly speaking, the best *thing* that exists; existence is not a thing. Existence, rather, is a phenomenon or, rather, *the* phenomenon. If we are to be truly accurate, we would have to say not “Existence is the best existing thing” but “Existence is the best existence.” As I have stated earlier, existence may be the best existence, but that tells us nothing; what we want to know is what are the best things or what is the best thing in existence. If, for instance, existence is good because it is necessary for goodness, which truly is, by definition, the best thing, *because it is a thing*, then anything that is better for goodness, it would seem, is better. This discovery of better things for goodness, however, necessitates our picking out different things in existence. In other words, while existence is good, there are entities in existence that must be better than mere existence; there must be entities in existence that are better than each other due to their comparatively greater tendential necessity for tendential necessity. Thus, we must recognize what we are talking about as a linguistic-conceptual framework: The Existence Framework. The Existence Framework, we have acknowledged, is itself intrinsically good, but we must realize that there are other frameworks within the Existence Framework that might be better than existence. How could something be better than existence? We shall shortly see exactly how this is the case, but, for now, we can think of it in the following way: If a framework, itself, is intrinsically good, then if there is another framework within that framework, then that framework is also good. The framework-within-the-framework, however, might have the potential to be better than the initial framework because the framework-within-the-framework *literally* is the initial framework *plus* some additional framework. Thus, the

framework-within-the-framework can be better than the initial framework because it represents an addition to, and not a negation or replacement of, the initial framework.

Now it is time to clarify an assertion I offered above. I stated that “there must be things in existence that are better than each other due to their comparatively greater tendential necessity for tendential necessity.” This, however, is not entirely accurate. In order to have something be better than something else, it is not the case that the thing in question must be better than something else due to its comparatively greater tendential necessity for *merely* tendential necessity; that would make it merely better for goodness and not *better* or, as it were, good for what is better or better for what is better. Again, unlike most or all other properties, goodness is reflexively intrinsic. As was the case with goodness, itself, we had to determine what was intrinsically good by self-reference, self-entailment, and tautology. We had to do what I called self-proof. As we have also previously seen, this necessity for self-proof is due to the peculiar nature of goodness as goodness-for, this peculiar nature is due to the peculiarity of tendential necessity essential to goodness, and this, in turn, is the result of the concept of necessity itself. We see such a relationship to necessity in almost no, if any, other properties. We cannot, therefore, as with other properties, simply engage in comparison in the usual way. We cannot, for instance, take the tendential necessity of the tendential necessity of functions, by itself, and ask “What is more or most tendentially necessary for the tendential necessity of functions?” Why is this? This is because tendential necessity has no true object besides goodness or tendential necessity as its function; it does not, as of yet, have what is *better* as its object. We are not asking any questions, in other words, about what is a more tendentially necessary planet, or hat, or crop. Planets and hats and crops, strictly speaking, are not necessary for the tendential necessity of functions, although they are, indeed, *tendentially necessary* for it; they are not necessary for goodness. They are, instead, sufficient for it. This is why I have previously stated that there is no standard of comparison inherent in the concept of goodness itself, at least not a non-vacuous one.

But the problem remains: we still have to choose, and we have to choose what we judge is better; that is what we do when we act voluntarily, and that is what we do when we make any judgment or think about anything. All of this activity—all of these functions—presuppose a better choice. This better choice, furthermore, is often *not* with reference to “better” as “better-for some specific function X (such as being a thief)” although, much of the time, it clearly is. We cannot simply gloss over this fact; we cannot simply pretend it does not exist. In fact, we literally cannot pretend this because even pretending this presupposes a judgment that it is better to pretend that we do not have to make judgments concerning what is better. To judge that nothing is better than anything else, however, entails a contradiction, since to judge that nothing is better than anything else is to assert that it is better to judge that nothing is better than anything else. Foot tells us that the grammar of goodness is meaningful only within the Life Framework, and I believe I have shown this to be false. While this does not discount the fact that “good” is still meaningful within the Life Framework, she and I are stuck with the same problem: the reflexive intrinsicity of goodness prevents there being a clear substantive standard for

comparison of goodness because, as it is founded upon the concept of necessity, it limits judgments of what is better to what is better for goodness, which severely limits us to what is better or best for existence itself or, on Foot's account, life itself.

There is, however, such a property as betterness, and it is founded, as it should be, upon the property of goodness. Betterness, however, has the additional element of comparison added to it that is inherent in all extrinsic properties. If goodness, itself, is not inherently comparative in the same way that other properties are, then betterness is, and this makes sense. One might object at this whole project and state "What is this business with betterness? It seems as though you are simply trying to graft a concept onto a theory where it does not fit. Why can't you simply be happy with the concept of goodness and move on? Why do you need to graft this concept of betterness onto your conceptual framework?" These are good questions, and I will take them up now. To start, I am not simply trying to graft the concept of betterness onto this theory or force-feed the reader this concept of betterness. On the contrary, I am going about establishing meta-ethical theory the way any theory, meta-ethical or not, should be established, and that is by providing a framework that has the most explanatory force. What does explanatory force mean in this context? Explanatory force, here, refers to the phenomenon of value judgment itself, let alone choice. A theory which explained and provided a framework for the good and left out the better would be a very deficient one indeed. Why is this? Is it because I have a cultural bias toward thinking that some things are better than others? No. On the contrary, it has been clearly demonstrated anthropologically, neurophysiologically, psychologically, linguistically, and philosophically that all agents make value judgments concerning what is better or less worse and not simply concerning what is good. Since this phenomenon is a very real one—more real to us and our essential natures than perhaps anything else—we must explain it. Certainly, we can explain it, from certain perspectives, including neurophysiology and psychology, but these prove inadequate. Why? They prove inadequate because they do not determine for us to what we are referring or what we are really saying when we make judgments concerning what is better; they do not give us an account of why we say that some things are better than others or what we actually *mean and logically imply* when we assert that some things are better than others. If we are going to explain this phenomenon, we will, therefore, have to understand what we *mean and logically imply* when we say such things, and that is precisely what I intend to do.

6.1.10 *Objection to Betterness and Reply*

An objection might arise, namely that there is, indeed, a non-vacuous standard of comparison that is built in, as it were, to the concept of goodness itself. My account of "good for a thief," the objector might assert, is proof of this. For instance, we can pick any essence and *teloi* we want and figure out, simply from the concept of goodness, what is good, better, and best for these respective functions. We can pick

anything, she might assert, including such things as cars, boats, flowers, birds, rivers, planets, supernovae, and even people. And if we can do all this, she might ask, why do we need a separate concept of betterness? The answer to this is simple, and I will answer it with a question: Which of these essences and *teloi*, pray tell, should we or would it be better for us to promote? People, we might judge, are worthless; they are worse than, say, flowers. Flowers, however, are better yet than supernovae, but supernovae are better than rivers. Do you see the problem now? The problem, to be precise, is that, although we can pick out some random essence and *teloi* and judge the comparative goodness of various functions in relation to its *telos* or, to be clearer, the comparative tendential necessary of various functions for its *telos*, we have, as of yet, no method for determining exactly which essences and *teloi* we are to judge as better in the first place; we have no coherence in our judgments. All it seems we can do is simply “pull an essence and *telos* out of a hat,” use it as the standard against which we judge everything else, and then go full steam ahead in making judgments. We can pick as many essences and *teloi* out of as many hats as we want; we will all still be wrong in our assertions of which essence and *telos* it is better to frame our judgments against, or, if it is not the case that we are all wrong, we will all at least be unjustified in doing so because we will never have any non-arbitrary, justifying reason to pick out any specific essence or *telos*. It seems clear, therefore, that it is not possible, whichever method we have tried thus far, to derive a judgment of what is better from a judgment of what is good. To reiterate and to keep the reader on track with my argument, this impossibility of derivation of what is better merely from what is good is the result of two basic features of goodness: (1) Goodness has peculiarities founded upon the concept of necessity that make it reflexively intrinsic. This self-referential feature makes it, by itself, impossible to compare to anything but itself if not applied to some specific essence and *telos* and; (2) When goodness is applied to some specific essence and *telos*, the comparative goodness of various functions can be measured against each other only in relation to that specific essence and *telos*. As a result, however, the specific essence and *telos* chosen are arbitrary and cannot be determined, via goodness itself alone, by some overarching standard. I hope, by now, that the reader has seen why this concept of betterness is not simply a whim or bias of mine; it is necessary to analyze if we are ever to know what we are saying or talking about.

6.1.11 A Naturalistic Definition of Betterness

What, then, do we actually mean when we say that something is better than something else? Language, as we know, has a firm basis in logic. Yes, it is true that language is also often, and importantly, used to express emotions, psychological states, and qualia in general, but language, and especially deep grammar, takes its basic structure from logic. What, then, is the logic of betterness? As I stated before, if we are to discover what betterness is, we must ground our understanding of betterness in our understanding of goodness. Goodness, we have shown, is the

tendential necessity of functions. When we state that something is good for something else, we state, thus, that some function *X* is tendentially necessary for some function *Y*. When we want to state that something is better for something else, however, we have to make a comparison between tendential necessities. We must state, in other words, that some function *X* is more tendentially necessary or has greater tendential necessity or has a greater tendency to lead to or yield *Y* than some other function *Z* does. *X*, therefore, is better for *Y* than *Z*. The problem, as I previously stated, however, is that we cannot actually show *Y* to be better in itself; we cannot, thus, judge that *Y* ought to be or categorically must be the case in the first place. In addition, due to the reflexive nature of goodness, if we try to use goodness itself as our standard for demonstrating *Y* to be better, we will fail because we can only show *Y* to be good and not better than anything else. What, then, does it mean for something to be better? We know that, in order for something to be better or best, we need to have a standard for comparison. This, in fact, is the case for the commonest properties such as width, largeness, and mass; in order for us to say that something is wider than something else, we need something that actually has more width with which to compare it. Likewise, in order for something to be better than something else, we need something that actually has more goodness with which to compare it. In the case of width, we need something that, in other words, has *comparatively* more width. In the case of goodness, then, we need something that has *comparatively* more goodness. Unlike for these other intrinsic properties, which are not reflexively intrinsic, the concept of comparison, and not simply the nature of the intrinsic properties themselves, stands at the center of what it means for something to be better and not simply good.

What shall we admit that betterness is, then? To state that something is better, we should see by now, is to state that something is comparatively more tendentially necessary. But comparatively more tendentially necessary for what? Certainly we will not say that something is comparatively more tendentially necessary for the tendential necessity of functions, for we will be stuck right back in the same condition; we will simply be saying that something is better for goodness than something else or, in other words, that existence is better than non-existence for goodness or that goodness is better for goodness than non-goodness. This, again, does not solve our issue. We do not want to know what is wider for wideness, what is larger for largeness, or what is more massive for massiveness. We wanted to know, rather, what is wider in comparison to the wider or widest thing, what is larger in comparison to the larger or largest thing, and what is more massive in comparison to the more or most massive thing. Thus it is with goodness. We did not want to know what is better for goodness. We wanted, rather, to know what is better in comparison to the better or best thing. What, then, is simply better or best? To figure this out, we will, once again, have to revisit the concept of reflexive intrinsicity. In order for some function to be better, we recall, it is necessary that that function be comparatively more tendentially necessary. But comparatively more tendentially necessary for what and, most importantly, compared to what? The clue is in our immediately prior statement. Let us consider it again, since it bears repeating: "In order for some function to be better, we recall, it is necessary that that function be comparatively more tendentially necessary." This is

important, for re-written in another way, we see that we are truly stating, “It is a necessary condition of a function X’s being better that it be comparatively more tendentially necessary” or, more clearly, “Function X’s being comparatively more tendentially necessary *is necessary for* function X’s being comparatively more tendentially necessary.” Although a tautology, it is an important one since, as I cannot emphasize enough, the definition of goodness is so inextricably bound up with the concept of necessity. Remember that, if some function X is necessary for some function Y, then function X is better than not-function-X for function Y. In other words, X is better than not-X for Y. If X and Y are identical, which they are in the statement “Function X’s being comparatively more tendentially necessary *is necessary for* function X’s being comparatively more tendentially necessary,” then it follows that function X is better than not-function-X for function X. We can conclude, in fact, that function X is the best for function X than any other function could possibly be. We can see, then, that betterness, itself, is better and even the best for betterness.

But what, still, makes betterness itself? What is essential to betterness or, if you like, comparatively greater tendential necessity, that makes it betterness and not some other thing? We know that what is essential to goodness is necessity, tendency, and function or, in other words, the tendential necessity of functions. We know it is good, thus, for there to be existence, for there to be functions, for there to be necessity, and for there to be tendency. In other words, we know it is good for there to be goodness and that goodness is good. As with goodness, therefore, necessity is better than non-necessity for betterness or the comparatively greater tendential necessity of functions. I hope this is clear. Tendency, likewise, is better than non-tendency for betterness or the comparatively greater tendential necessity of functions. I hope this is also clear. I hope, also, that it is clear that function is better than non-function for betterness or the comparatively greater tendential necessity of functions. All of these things, however, are essential to goodness and not, strictly speaking, to betterness. We see, then, that there is only one other possible concept that could be counted as necessary and, therefore, better and, certainly, even *best* for betterness or the comparatively greater tendential necessity of functions. That concept is comparativeness itself. Comparativeness, we can safely say, is not only better than non-comparativeness for betterness. It is, rather, what is *best* for betterness since it is categorically necessary for betterness; it, along with the other concepts comprising goodness, constitutes the essence of betterness or that without which betterness would not be betterness or, if you like, that without which the concept of something’s being better would be impossible. This is something very interesting indeed, for this means that we can say, with impunity, that comparativeness is *intrinsically better* than all of those things that are not comparativeness. Alternatively, we can state that those things which are comparative are intrinsically better than all of those things that are not comparative, since something’s being comparative is a necessary and, thus, the most tendentially necessary function for comparatively greater tendential necessity of function; the emphasis is on the fact that something’s being comparative constitutes its being a tendentially necessary function itself, and, in addition, the most tendentially necessary function for the comparative tendential necessity of functions. The above

statements are tautological and, thus, necessarily true. We can see this more clearly in the following statement: “Comparativeness is the most tendentially necessary function for the comparatively greater tendential necessity of functions.” We can break this down even further to make the tautology clearer: “Comparativeness is the function that has the greatest tendency to yield the comparatively greater tendency to yield functions” or “Comparativeness has the greatest tendency to yield the comparatively greater tendency to yield. . .” or “Comparativeness has the greater tendency to yield the comparatively greater tendency to yield. . .” Since we know from previous analysis that it is a necessary condition of something X’s tendential necessity for something else Y that it at least not be impossible for X to lead to or yield Y, and since we also know that comparativeness is a necessary condition of something X’s being comparatively more tendentially necessary for something Y, we can form the statement “Comparativeness is necessary for comparatively greater tendential necessity. . .” and form, from this, “Comparativeness is better for comparatively greater tendential necessity (than non-comparativeness)” and “Comparativeness is the best for comparatively greater tendential necessity.” Consequently, we can state “Comparativeness is better for betterness (than non-comparativeness)” and “Comparativeness is the best for betterness.” We can conclude, thus, “Comparativeness is better (than non-comparativeness).” We can follow a relevantly similar process for bestness simply by replacing every instance of the expressions “comparatively more tendential necessity” and “comparatively greater tendential necessity” with “comparatively the most tendential necessity” and “comparatively the greatest tendential necessity,” respectively. We thus yield the statement, “Comparativeness is the most tendentially necessary (of all other functions) for the comparatively most tendential necessity. . .” From a similar method of deduction, we eventually end up with the conclusion: “Comparativeness is the best (of all other functions)” and, thus, “Comparativeness is the best thing that exists.” This is a huge, monumental claim, and that is why I thought it necessary to spell out the tautology in several different ways. Comparativeness, I should clarify, I am specifically defining as that property which is essential to the class of things which are comparative *and* tendentially necessary for something, which will become important momentarily. Thus, we should more precisely state “Things which are comparative *and* tendentially necessary for something are necessary for comparatively greater tendential necessity” and “Things which are comparative *and* tendentially necessary for something are better than things that are non-comparative *and* non-tendentially-necessary-for-something.”

6.1.12 *Objections and Replies*

Let us deal with some possible objections at this stage. One might, first, make the objection “This is all a game of semantics; we do not need the superfluous term ‘comparative.’ If we simply reformulate ‘comparatively greater tendential necessity’ and ‘the comparatively greatest tendential necessity’ to ‘more tendential necessity’ and ‘the most tendential necessity,’ respectively, all we end up with is

‘more tendential necessity is necessary, and thus the most tendentially necessary, for more tendential necessity’ and ‘the most tendential necessity is necessary, and thus the most tendentially necessary, for the most tendential necessity.’ We do not need to make reference to comparativeness at all.” Let us answer this. “Comparativeness,” we must admit, is not simply a superfluous term here. Even if we examine the expressions “more tendential necessity” and “most tendential necessity,” what is essential to these expressions is not tendential necessity, alone, but tendential necessity *in addition to* comparison or a comparative relationship that something has to something else; “more” and “most” are comparative terms in that they pick out a relationship between something that has a greater quantity of the property in question and something that has a lesser quantity of the property in question. With width, for instance, when we state “The table is wider than the desk,” there is a concept, here, that is essential to this statement without which the statement would not be itself and, furthermore, could not exist. If we rephrase it, then, by explicitly acknowledging this element, we end up with such statements as “The table is comparatively more wide than the desk” or “The table has comparatively more intrinsic width than the desk.” Thus it is similar, although not the same, with goodness. “The table,” we would have to admit, “is comparatively more tendentially necessary than the desk” or “The table has comparatively more tendential necessity than the desk.” As we have stated often, however, tendential necessity does not stand alone; tendential necessity is tendential necessity *for things*. Thus, in order to state that something is better, we would not state that something has comparatively more tendential necessity, but that something has comparatively more tendential necessity for something’s being better. The something that has such comparatively greater tendential necessity is comparatively greater tendential necessity itself. To see why it is the term “comparativeness” that we are picking out specifically, and not specifically the expression “the greatest comparative tendential necessity,” let us consider something I highlighted above, namely that comparativeness *is* a tendentially necessary function or is, otherwise, something that is tendentially necessary for something. Thus, while it is the case that comparativeness is a function and that it has tendential necessity, it is only relevant to something’s being comparatively more tendentially necessary that something be a function that has the greatest tendential necessity to be comparative, and that is comparativeness. This bears repeating: *that function that has the comparatively greatest tendential necessity to be comparative is comparativeness itself*, although, as we shall see, relationships, themselves, can have comparativeness or be comparative in relation to something else. Thus, when we state “Comparative tendential necessity is comparatively most tendentially necessary for comparative tendential necessity,” we are simply stating “Comparativeness, which is a function that has tendential necessity, is a function that has the most tendential necessity for the comparatively most tendential necessity” or, even more simply, “Comparativeness is necessary for something to be comparatively more tendentially necessary.”

Let me now clarify something. The expressions “being comparative” and “having comparativeness” are not used here to imply observer relativity. Why is this?

Consider that *widerness*, if you will, necessitates, of course, width. Something is *wider* than something else, however, without someone's comparing it to be wider than that thing. In other words, it does not matter if someone is comparing the thing in question; this *widerness* is not, then, dependent upon comparison itself for its being wider. I am not speaking of comparison in this observer-relative way, in that, in order for something to be comparative or have comparison, there must be an observer to compare the things in question.

6.1.13 Teleological Comparativeness

This brings us to the next objection. "Why," one might ask, "is comparativeness or something's being comparative necessary for something's being better but not necessary for something's being wide? Both terms are predicated on the use of comparison, so why is comparison not necessary for *widerness*? Is there not a discrepancy here?" There is, indeed, a discrepancy here, but one that is perfectly sensible. Something's being comparative or having comparativeness, we said, is not necessary for something's being wider, but only in the observer-relative sense. In order for X to be wider than something else, both objects must have comparison and be comparative, but in a specific way. They must be comparative in relation to each other; again, this is not to say that they must *be compared to each other*, implying that there is some observer or comparer that is necessarily involved in something's being wider. Indeed, X can be wider or less wide than Y if no observers or comparers exist. We must remember, however, the special peculiarities of goodness and betterness. For such properties as width and mass, in order to say that something is wider or more massive, we imply that those things have more width or more mass than some other things. For goodness, however, we cannot simply state that those things are better and the best, respectively, which have greater and the greatest tendential necessity. Why is this? Yet again, this is because of the feature of tendential necessity inherent in goodness and, therefore, betterness as well. If we are going to state that something has tendential necessity, we need to pick out that *for which* it is tendentially necessary. If we were, for instance, to look around for things in the universe which have tendential necessity and state "Look there—that rock has more tendential necessity than this sulfuric acid" or "This mercuric oxide has more tendential necessity than this book," we would be, rightly, very confused. This, I am arguing, is not how we speak about something's being better; this is not how we use the term "better." Tendential necessity, we must emphasize once again, always makes reference to a *telos* through the addition of the word "for;" *this teleological way, indeed, is the only way it can be meaningfully used*. Again, then, in order to truly understand what we are talking about when we say that something is better than something else, we must go to the very heart of what it means for something to be better; something's being better, and not simply something's being better-for, is ultimately dependent on the degree to which various functions are tendentially necessary for this *telos*, betterness. We might even go so far as to speak

of a Betterness Framework, if this is a helpful concept. Thus, unlike most other intrinsic properties, goodness and betterness require *teloi* for comparison, and this requirement is built into the Betterness Framework. Thus, *widerness* and *betterness* both require comparison or require being comparative, but *widerness* requires quantitative comparison, whereas *betterness* requires, first, teleological comparison and, then, quantitative comparison. This teleological reason, thus, does not exclude such comparative properties as width, depth, and largeness from being comparatively more tendentially necessary (than not) for comparativeness. Instead, it simply demonstrates that such properties as width, depth, and largeness are not essentially comparatively tendentially necessary for comparativeness *of some function X to some function Y with regards to a telos*. In other words, although it might sometimes be the case that we refer to width-for, depth-for, and largeness-for, we do not usually use these terms in this way. Of course, if they are used in this way, then they simply become part of the teleologically comparative domain of discourse of which we are speaking anyway. This is an important distinction. Does it help us find something that is comparatively more tendentially necessary for betterness than these properties? I believe so.

This distinction leads us to our next observation, namely that, while everything that has an essence has a *telos*, it is not possible for just anything that has an essence to be its own *telos*, or, if you will, for the *telos* to be reflexive. Thus, when we state that X is wider than Y, it is not the case that we are claiming that X is wider than Y because it is more comparative than Y or because it has more comparativeness than Y, teleological or otherwise. We are, instead, stating that X is wider than Y because it has *comparatively* more width than Y; we are not, in other words, claiming that X is a type of comparativeness that Y is not or that *widerness* is a type of comparativeness. We are, however, claiming that this is the case with betterness. Because of the teleological nature of goodness and betterness, in order to state that X is better than Y, and not simply that X is better than Y for Z, *we must find the telos of betterness*. Let us recall the essence of betterness. It is, as we explained, the comparatively greater tendential necessity of functions. But what is its *telos*? The statement above is instructive: “While everything that has an essence has a *telos*, it is not possible for just anything that has an essence to be its own *telos*, or, if you will, for the *telos* to be reflexive.” Betterness is reflexive, but why, more specifically, is it reflexive? To answer this, we must remember that goodness and betterness concern tendential necessities of functions. Consider, now, that comparativeness or something’s being comparative is a function. Consider that comparativeness or something’s being comparative is that which makes the comparative tendential necessity of functions possible. Consider, also, that comparativeness is necessary and, thus, tendentially necessary for itself. Comparativeness or something’s being comparative, thus, must be that function which is most tendentially necessary for the *comparative* tendential necessity of functions and not simply for the *tendential necessity of functions*. But now, we should see that we can take this a step further. It is not enough simply to state that comparative tendential necessity necessitates comparativeness. We must narrow our search further. What kind of comparativeness is necessary for comparative tendential necessity? The answer, we should see, is

teleological comparativeness or, more specifically, the comparativeness of tendential necessities to each other with regards to a specific *telos*. Now, therefore, we must go on a search to find out which thing or things in the universe possess teleological comparativeness or which thing or things in the universe possess or are comprised of teleologically comparative relationships.

6.1.14 *Objections and a Reply via the Argument to the Commensurability of Teloi*

One might object that comparativeness or something's being comparative is that which also makes something's being wider possible, but this is irrelevant. As we have previously stated, X's being wider than Y depends on *the quantity of X's width in comparison to Y*. Width, here, is obviously *not* comparativeness or something's being comparative. By contrast, X's being better than Y depends on *the teleology of X and Y and the quantity of X's goodness in comparison to Y relative to some telos Z*. It is this teleology that distinguishes these two types of properties. Here, one might state, goodness is obviously not comparativeness or something's being comparative, and one would be right. However, one would be ignoring the teleological component. "What," we must ask, "is the teleology?" Whatever the teleology is (it could be survival, or stealing successfully, or anything else), it is the standard of X's *being better* than Y *for Z*. If the teleology, however, is not something so particular but is, say, *something's being better and not simply better-for*, then we must ask what the essence of betterness, itself, is. Betterness, in fact, is actually a *type of comparativeness*, namely the comparative tendential necessity of functions. But how is this, strictly speaking, a type of comparativeness? I did not, for instance, give the expression "the comparing of tendential necessities;" instead, I gave the expression "the comparative tendential necessity of functions." Are they the same thing? First, the answer is no; they are not the same thing; the first is an action which entails the possession of the property of comparativeness, while the second is a property which does not necessarily entail the action of comparing. Consider that everything has a tendential necessity for some function. Keep in mind, also, that the greater or lesser tendential necessity for these different functions is independent of observers; the relationships between these phenomena do not necessitate comparison or being compared in order to be comparative or, otherwise, to have the relationships of tendential necessity that they actually have to each other. Since all things have *teloi*, however, these *teloi* must have comparatively tendentially necessary relationships to each other relative to some other *teloi*, and so on. In other words, there is some necessary condition of comparativeness that makes such teleological comparativeness possible and intelligible in the first place, and that is commensurability of *teloi* or, if you like, the fact that some *teloi* must be better than others and, ultimately, that some *telos* must be the best.

Consider what we stated earlier, namely that we can pick out anything, including cars, boats, flowers, birds, rivers, planets, supernovae, and people, and, in each individual case, find the respective essence and *telos* of each, and, then, compare the comparative tendential necessity of various functions against each other and rank them, relative to the *telos* in question, according to how comparatively tendentially necessary each function is for that *telos*. This, as we said, is all good and well, but, as I also asked above, which of these essences and *teloi* are actually better than the others? I stated: “People, we might judge, are worthless; they are worse than, say, flowers. Flowers, however, are better yet than supernovae, but supernovae are better than rivers. . . . The problem, to be precise, is that, although we can pick out some random essence and *teloi* and judge the comparative goodness of various functions in relation to its *telos* or, to be clearer, the comparative tendential necessary of various functions for its *telos*, we have, as of yet, no method for determining exactly which essences and *teloi* we are to judge as better in the first place; we have no coherence in our judgments.” Thus, we seem to have a partial solution, that is, we seem to have found the essence of betterness itself, which is the comparative tendential necessity of functions, and we seem to have found the essential necessary condition of comparative tendential necessity of functions, which we said was comparativeness, but we seem to be lacking something. Why does it seem so? Because we are. What, exactly, are we lacking? Consider that comparativeness is a very far-reaching concept and that “things that are comparative” is a very extensive class of things. The question now is “Which thing that is comparative is better for betterness or is comparatively most tendentially necessary for the comparatively greater tendential necessity of functions?” There is a multitude of comparative things which could fill this gap of serving as the standard against which we judge everything to be better. Which comparative thing enables and makes coherent statements concerning whether rivers are better than rocks or whether supernovae are better than flowers? Which comparative thing, in other words, gives such statements coherence? It seems we need to find a more specific *telos* than simply “things which are comparative.” Teleological comparativeness, although still very general, narrows our search.

Consider that, within the Goodness Framework, the concepts relevant to goodness are limited to things that have functions and tendential necessities. Such a question as “Why is function good?” is, therefore, needless; it is good simply because it comprises the essence of goodness, namely the tendential necessity of *functions*. This is like asking, “Why are animals living things?” and rightly receiving the answer, “Because the concept of life is inherent in the concept of an animal because it is meaningful only in the Life Framework.” Thus it is, too, with betterness. To ask such a question as “Why is teleological comparativeness or something’s being teleologically comparative better than non-teleological comparison or something’s being non-teleologically comparative?” or “Why is teleological comparativeness teleologically comparatively more tendentially necessary for teleologically comparative tendential necessity than non-teleological comparison?” is to prepare oneself for the answer “Because teleological comparativeness or something’s being teleologically comparative is essential to the concept of what it means for something to be

better because it is meaningful only in the Betterness Framework.” If we acknowledge this, then the statement “Teleological comparativeness or something’s being teleologically comparative is better than non-teleological comparativeness or something’s being non-teleologically comparative” becomes obviously true. If we admit that “Teleological comparativeness or something’s being teleologically comparative is better than non-teleological comparativeness or something’s being non-teleologically comparative,” we will have to accept the statement “Something’s being teleologically comparative is teleologically comparatively more tendentially necessary for teleologically comparative tendential necessity than something’s being non-teleologically comparative.” What things, in general, can we say are comparative? Certainly, relationships between properties are comparative, such as those that comprise width, depth, and largeness, but that does not give us any help, namely because we have no standard which tells us which relationship to pick out. Relevantly similar to what we asked above, exactly which relationships between properties are better or best? To state simply that relationships between properties are comparative does not yield which properties are better and certainly not which specific relationship between properties is best. So then which specific relationships between which specific properties are comparatively more tendentially necessary for comparative tendential necessity than which specific others, and, moreover, which specific relationship between which specific properties is the most tendentially necessary for comparatively greater or greatest tendential necessity? We narrowed our search by considering teleological comparativeness, but that, by itself, renders little information. Let us, then, dig deeper. We must ask the following question: “Is there more than one thing in the universe that has teleological comparativeness?” If the answer is “yes,” then we know we have not narrowed down our search enough, for we will still not know which thing with teleological comparativeness we need to compare different tendential necessities to; there is no *telos*. If the answer is “no,” then we have found our *telos*. We have reason, I shall argue, to conclude that there is only one thing in the universe that possesses the property of teleological comparativeness.

6.1.15 Comparison and the Telos of Betterness

The answer to our question, we must see, is the act of teleological comparison of tendential necessities. What I mean by “the act of teleological comparison of tendential necessities” is the process by which something is judged to be more or less or most or least tendentially necessary for something or, in other words, the process by which something judges something to be better or worse or best or worst. I am purposely not using the simple expression “the act of comparison” because this is ambiguous. To compare, which entails the act of comparison, could entail to compare what is wider or widest. This, however, is irrelevant to the act of comparison *as a tendential necessity or comparatively greater tendential necessity* for something. Yes, comparison, like all things, is inherently tendentially necessary for something else, but “comparison,” as we use it with regards to the expression “the

teleologically comparative tendential necessity of functions,” *does not* denote such non-teleological comparison; in other words, when we are trying to find out what is better or best for betterness, those types of comparison that do not denote tendential necessity (such as “wider,” “more massive,” or “lighter”) or, otherwise, which do not denote the comparison of something to something else relative to a *telos* (since inherent in the idea of tendential necessity is the idea of *telos*), are not, by definition, being used to denote tendential necessity for anything; these types of properties are therefore irrelevant. One could, of course, use such adjectives as “wider,” “more massive,” and “lighter” in a way that denotes tendential necessity for something, as in the expression “X is wider than Y for Z,” but, once again, these adjectives then become one of the many adjectives that are on our smorgasbord of comparatively tendentially necessary functions from which to choose in order to discover which comparative thing, among comparative things, are most tendentially necessary for teleologically comparative tendential necessity.

We have, in this process, actually narrowed down which comparative things, among comparative things, can be most tendentially necessary for comparative tendential necessity. We have narrowed down that it is things which are comparative *and* also tendentially necessary, and not things which are simply comparative, that can be most tendentially necessary for comparative tendential necessity. This might seem obvious to some, but it was again necessary to elucidate this both for clarification and to fend off any possible objections pointing to the irrelevant fact that adjectives other than those that denote tendential necessity are comparative.

Let us more directly address what comparative thing it is that is comparatively most tendentially necessary for comparative tendential necessity. It is, once again, the act, itself, of the teleological comparison of tendential necessities. Let us ponder this for a bit: Why would the act, itself, of teleological comparison of tendential necessities have the greatest tendency of all other things to yield teleologically comparative tendential necessity or teleologically comparatively more tendential necessity? Consider that, before, we admitted that, in order to have any statement be true about what is better (birds, rivers, flowers, supernovae, etc.), and not simply about what is better-for, it would be necessary for the truth of any of these statements that betterness have a *telos* that constitutes a standard against which anything can be shown to be better or worse than anything else. In other words, such a *telos* of betterness is a necessary condition of the truth of such statements concerning betterness. Teleological comparativeness, we have said, is the best for this *telos*, which is teleologically comparative tendential necessity, but we have said that we have not discovered exactly what kind of teleological comparativeness or teleologically comparative thing is best. Let us, then, consider the following: Without the act, itself, of teleological comparison of tendential necessities, the *telos* of betterness, that is the comparative tendential necessity of different *teloi* and not simply the comparative tendential necessity of different functions relative to a *given telos*, would be impossible. Why is this? As we have seen, there is nothing that can be derived from goodness, itself, which gives us betterness, and there is nothing else which we have been able to discover which could possibly make the betterness to which we all refer every day of our lives possible. It seems, thus, that it is the act of teleological comparison of tendential necessity, itself,

which yields betterness. One might once again object but, this time, state, “But we do not need such a teleology. Perhaps it is the case that we all simply judge that it is better to wake up in the morning because living one’s life out of bed is better than living one’s life in bed and, further, that living one’s life out of bed is better because life is better. That life is better than non-life might be all we need.” This, however, is not all we need. For instance, we might still ask “But why is life better than non-life?” Pick some other arbitrarily judged thing, say friendship, eating, or music. If one chooses one of these things, one might judge music, for instance, to be better than non-music. “But why?” we could rightly ask. “Why is music better, and what is it better for?” Some response A might be given. Music is, then, better for A than eating is. But then why is A better than not-A? For B, of course. But then why is B better than not-B? For C, of course. But then why is C better than not-C? It is clear that this process goes on *ad infinitum*. This kind of judgment, however, does not capture the kind of judgments we actually inherently make concerning what is better or teleologically comparatively more tendentially necessary. When we make judgments concerning what is a better state of affairs or a better thing, while we sometimes follow the aforementioned pattern of judgment, we are always, at least implicitly if not explicitly, ultimately uttering a statement about what is better and not simply about what is better-for; the process of judging something to be better-for cannot, as we have seen, go on *ad infinitum*; somewhere at the end of our process of judging, we have to make a judgment concerning what is simply better; the process, in other words, must stop at some point. When we say, for instance, that not-killing is better than killing, we are not simply stating that killing is worse for life or that not-killing is better for life; we are stating that not-killing is simply better. When we have lunch with our best friend, and we discuss something of importance to us, it is not simply better for us or better for our psychologies, we say; instead, we claim that it is simply better, better that is, to eat lunch with our best friend and discuss something of importance. As I have stated earlier: “We do not want to know what is wider for wideness, what is larger for largeness, or what is more massive for massiveness. We wanted to know, rather, what is wider *in comparison to the wider or widest thing*, what is larger *in comparison to the larger or largest thing*, and what is more massive *in comparison to the more or most massive thing*. Thus it is with goodness. We did not want to know what is better for goodness. We wanted, rather, to know what is better *in comparison to the better or best thing*.” Thus, in order to make any statement concerning what is better, we must make some claim concerning what is the better or best thing or, alternatively, what is teleologically comparatively more or most tendentially necessary, and thus implicitly what is teleologically comparatively more or most tendentially necessary *for* teleologically comparative tendential necessity, since it itself is necessary for the very concept of something’s being better; *an implicit citation of its betterness and indeed bestness is a necessary condition of claiming that anything else at all is better or best*. The question is, “Can any such claim of one thing’s betterness than another thing be substantiated?” The answer is “yes.” If the answer were “no,” we would have to admit that all of our judgments concerning what is better are groundless and that, thus, there is no reason to do anything. This, however, we have shown to be false in Chap. 3. We have shown,

namely, that there is reason to do things that do not entail invalid judgments; this, alone, demonstrates that we have at least some reason to do something. But we have more than *some* reason to do *something*; in fact, we have strong reasons to do certain determinate things, as is shown by Gewirth and Habermas's respective theories.¹¹ Thus, it seems that we would be remiss to overlook all of this and simply state, "We have no reason to do anything, and there is nothing we can do that is better or judge to be better than anything else." To state this is to completely overlook all of the analysis in Chap. 3. Thus, we should acknowledge, lest some nihilists or relativists make the charge to the contrary, that we have not come to the conclusion that there is something objectively better *because* we have exhausted all other options. Rather, if we take Chap. 3, *as well as* the present analysis into account, we can see that this conclusion is based upon our demonstration that there are reasons for doing and not doing certain determinate things *and* that the conclusion is also based upon our narrowing down those comparative things which could possibly be most tendentially necessary for comparative tendential necessity. Now, we have found those reasons and that teleologically comparatively tendentially necessary thing. It is because no other teleologically comparative thing is teleologically comparatively more tendentially necessary *for* teleologically comparative tendential necessity *than* the act, itself, of teleological comparison of tendential necessity that the act, itself, of teleological comparison of tendential necessity is teleologically most tendentially necessary for betterness. Why, once again, is this? It is because "things that are comparative" or even "things that are teleologically comparative" as classes, do not narrow down which thing that is teleologically comparative is *best* or *most tendentially necessary* for teleologically comparative tendential necessity; it simply tells us which class of things are so necessary. As a teleologically comparative thing, however, the act of teleological comparison of tendential necessity is narrow and specific enough to reveal what *specific thing* we can call better and, indeed, best. This, in turn, is due to the now revealed fact that, without the act of teleological comparison of tendential necessities, even the mere concept of teleologically comparatively greater tendential necessity, that is the concept of comparatively greater tendential necessity as a *telos* for all other *teloi*, would be impossible; the act of teleological comparison of tendential necessities truly is necessary, and thus the best, for betterness. Within the Betterness Framework, then, it seems an obvious conclusion to make that teleological comparison of tendential necessity is better and best for betterness than those things that are not teleological comparison of tendential necessity. In other words, the concept, itself, of teleologically comparing tendential necessities is necessary for the concept of teleologically comparative tendential necessity; nothing else, we have found, is so necessary because nothing else necessarily yields the phenomenon, itself, of teleologically comparative tendential necessity. Thus, it is indeed better to compare what is better.

¹¹ Although Habermas himself would not claim that his theory provides reasons to do "certain determinate things," on my reinterpretation of his theory, I have claimed that it does. See Chap. 3.

6.1.16 *Another Objection to the Commensurability of Teloi and a Reply*

An objection might be made at this point, namely that teleological comparativeness, and thus the act of teleological comparison, is not actually necessary for comparatively greater tendential necessity. One might state, instead, that all that is necessary for teleologically comparatively greater tendential necessity is a *telos* and at least two other things that can be compared to each other in relation to that *telos*. While it is correct that a *telos*, however random or arbitrary it might be, is necessary for teleologically comparatively greater tendential necessity because the existence of some *telos* is actually necessary for any teleological comparison of tendential necessity between any two objects in the first place, our objector has failed to address something important. When we ask what is teleologically comparatively most tendentially necessary for teleologically comparatively greater tendential necessity or what is necessary for teleologically comparatively greater tendential necessity (betterness, or the property of something's being better), we are not asking what thing is comparatively more tendentially necessary for *telos* X, *telos* Y, or *telos* Z. In other words, we are not asking what thing is comparatively more tendentially necessary for some random or arbitrary *telos*. We are also not asking what thing is necessary for some random *telos* X, *telos* Y, or *telos* Z; this, again, simply addresses *the class* of things which are necessary for teleologically comparative tendential necessity. Instead, we are asking *which specific telos* is more tendentially necessary than which other *teloi* for teleologically comparatively greater tendential necessity. Our objector might be persistent. "But any *telos* is necessary for teleologically comparatively greater tendential necessity because anything's teleologically comparative tendential necessity can be compared to anything else in relation to any *telos*." Again, this is certainly true, but let us do some re-analysis to find out where our objector has gone wrong. Teleological comparativeness itself, as a class of things, is teleologically comparatively more tendentially necessary than any other *telos* for teleologically comparatively greater tendential necessity. In other words, in order for any other thing to be compared to any other thing in relation to any other *telos*, there must be such a concept as teleological comparativeness. In other words, teleological comparativeness or "the property of being teleologically comparative" is that *telos* which is necessary for teleologically comparatively greater tendential necessity. But simply citing the class of things or *teloi* that are teleologically comparative tells us nothing; we must narrow our search. That *telos* within the class of "those things which are teleologically comparative" which is necessary for teleologically comparative tendential necessity is, we have seen, the act of teleological comparison because the act of teleological comparison is the *telos* that is necessary for the comparative tendential necessity of different *teloi*. With the vague concept of teleological comparativeness, we might be able to glean that something X is more tendentially necessary than something Y for the existence of a rock, but without the act of teleological comparison of tendential necessity, we can never answer the question of whether the existence of a rock is better than the non-existence of a rock. We might

be able to say that a rock is better than a non-rock for life, but we will never be able to say that life is better than non-life, and so on. Since we have seen that this sort of ultimate “better”-judgment contains a reference to betterness, which we have seen is defined as teleologically comparatively greater tendential necessity, we went on a search for what was teleologically comparatively the most tendentially necessary or, in fact, what was necessary for teleologically comparatively greater tendential necessity or, in other words, what was best for teleologically comparatively greater tendential necessity. We were not, in other words, on a search for what was best for what is teleologically comparatively more tendentially necessary *for*. Although, yes, we were technically on a search for what was teleologically most tendentially necessary *for* teleologically comparatively greater tendential necessity, it is that which is best *for* teleologically comparatively greater tendential necessity, or which is best *for* betterness, or that which is better that is, itself, simply better, where “simply better” and “simply best” can be taken to mean “better for what is better” and “best for what is best,” respectively. None of the functions that are better for *telos X*, *telos Y*, or *telos Z* are best *for* teleologically comparatively greater tendential necessity; instead, they are simply better or best for *telos X*, *telos Y*, or *telos Z*. To make this even clearer, although all of the functions that are necessary for *telos X*, *telos Y*, or *telos Z* are necessary for teleologically comparatively greater tendential necessity itself simply because they are comparatively tendentially necessary for *something*, neither X, Y, nor Z are necessary for teleologically comparative tendential necessity—they have only, thus far, shown themselves to be necessary for teleologically comparative tendential necessity *for*. What we were on a search for was that which is necessary for the concept, itself, of teleologically comparative tendential necessity or that which is better. If we found such criteria, we emphasized, we would have to take these criteria themselves to be simply better and best, for the betterness and bestness of such criteria are entailed by the betterness and bestness (instrumental or otherwise) of anything for anything else. Certainly, teleological comparativeness is necessary for that which is better, and certainly the comparativeness of two or more different things in relation to a given, arbitrary *telos* is necessary for the concept of betterness or what is better, but this is not what we were in search of; we were in search of what is necessary simply for teleologically comparatively greater tendential necessity and not the comparatively greater tendential necessity of X than Y *for* Z. If all we were doing was considering this concept of teleologically comparatively greater tendential necessity *for*, teleological comparison of the greater tendential necessity or greater value or betterness of something, in general, would be impossible; we would not be able to judge *anything* to be simply valuable; we would not be able to judge a flower to be more valuable than a rock, a planet to be more valuable than a nebula, or a person to be more valuable than a cloud of hydrogen. The point is that the act of teleological comparison of tendential necessity actually necessarily yields the concept of *teleologically comparative tendential necessity* and not *teleologically comparative tendential necessity for*. There is a vital difference here, and I want to make this as clear as possible. I am not saying, for instance, that the degree of tendential necessity of function X compared to function Y for the existence of a particular rock is dependent upon the act of teleological comparison of tendential

necessity; this relationship exists whether or not it is being observed. What I am saying, rather, is that the tendential necessity of this rock, being necessary for bacterial life, which is necessary for sentient life, which is necessary for further forms of life, and so on, eventually, at the end of all of this, requires the *act* of teleological comparison, and no longer simply comparativeness or even teleological comparativeness, in order to instantiate relationships of teleologically comparative tendential necessity of bacterial life, sentient life, and so on to each other in relation to a single *telos*, and not simply the relationships of X, Y, and Z to each other in relation to the *telos* of bacterial life, the *telos* of sentient life, or other various, random *teloi*. It is the existence of a single *telos*, and only a *single telos*, that makes the concept of teleologically comparative tendential necessity (or betterness), and not simply the concept of teleologically comparative tendential necessity *for* (or betterness *for*), possible. This single *telos* is the act of teleological comparison of tendential necessity. We see, thus, more clearly why such statements as “Not-the-act-of-comparison-of-tendential-necessity is better than the act of comparison of tendential necessity” is inherently contradictory; it is contradictory because it is tantamount to stating that that which is necessary for the concept of teleologically comparative tendential necessity is not tendentially necessary for the concept of teleologically comparative tendential necessity.

How does the act of teleological comparison make the concept of betterness possible? Consider that, when anyone performs any voluntary action, she does it because she believes that such an action is better than an alternative or set of alternatives; consider, also, someone’s assertion that X is better than Y. As we have previously stated, one might be making or implying instrumental assertions of betterness. It is impossible to make or imply any of these assertions, however, without simultaneously making or implying a statement of betterness and not simply betterness-for. Even in stating that it is better *for me* (or in my interest) to think or act in some way, for instance, I am endorsing these thoughts or actions and, thus, am simultaneously implying that it is *better* for me to think or act in this way or, in other words, that it is better to act in my interest. To explain it in deontic terms, I am claiming that, in order to fulfill standard E, I must (i.e., it is necessary or it is a necessary condition to) do X. Furthermore, I am making a claim about standard E, namely that I believe it to be something worth my valuing; in other words, it is not merely the case that I am assenting to the claim that I value it. I must, rather, assent to the claim that standard E is something that is worth my valuing; in other words, I must make a claim concerning standard E’s objective value to me. This makes my deontic judgments concerning standard E (i.e., I must do X), according to my own belief regarding it, categorically obligatory. I can also make objective judgments concerning the tendential necessity of different actions or other phenomena for standard E.

Again, our objector might tell us that we need not make such teleological judgments and that we need only pick some value out of a proverbial hat and act consistently with it. But this is literally impossible. We cannot pick out any random value out of a proverbial hat without making or at least assenting to some judgment concerning the objective betterness of that value or, in other words, without implicitly making the judgment “It is better that I value X.” One could even claim “It is better

for me to value X,” but one would still at least implicitly be claiming “It is better to value things that are better for me.” There is always an underlying claim of objective value here. Let us, however, indulge our objector. Let us suppose that it were possible to simply pick a value out of a proverbial hat and act in accordance with that value. In such a case, as it has been discussed earlier, none of our objector’s statements concerning his values could be shown to be any more valid than anyone else’s. Any and every action would, therefore, be permissible. Anyone could make any choices he wants, but none of his choices could be shown to be right ones, better ones, or remotely legitimate or justifiable in any way. All choices would be at once equally valid and equally invalid. The analysis of Chap. 3, however, has seriously undermined this idea; there must be something seriously flawed with theories which have such grounds as their bases. The idea, furthermore, flies in the face of everything we currently know about psychology, agentive behavior, and even language and logic themselves. Thus, what I mean when I say that the act of teleological comparison of tendential necessity is necessary for teleologically comparative tendential necessity is that agency, as Gewirth states, has a normative structure such that the concept of betterness, so-to-speak, comes with the package; agency, I mean, is inseparable from betterness. When a proposition is voluntarily uttered, when a word is voluntarily spoken, when a thought is voluntarily thought, the concept of betterness is invoked. It is in this way that, through their agency, agents discover this concept of betterness; their existence makes possible the expression and invocation of this concept in the universe. It is logically impossible for the state of affairs to be otherwise. We see more clearly now how the act of teleological comparison of tendential necessity is necessary for teleologically comparatively greater tendential necessity or the property of something’s being better.

6.1.17 Conceptual Connections to the Non-Descriptivist Fallacy

A counterargument to this account is highly problematic. If we were to claim that we refer to nothing at all by such statements and judgments, we will fall right back into what I said was a non-descriptivist fallacy: that, in our judgments of goodness and comparative goodness, we refer to absolutely nothing. Consider the following point: If function A’s bestness for betterness is dependent upon its necessity for betterness, then if function A is necessary for betterness, then function A is better than non-function-A. If function A is the act of teleological comparison of tendential necessities, therefore, then the act of teleological comparison of tendential necessities is better than not-the-act-of-teleological-comparison-of-tendential-necessities. It does not matter one bit whether or not one thinks that something’s being simply better than something else is an illegitimate relationship.¹² If we can prove that the act of teleological comparison of tendential

¹² What is an illegitimate relationship anyway other than one which is in some way disvaluable or which we must otherwise not accept?

necessities is necessary for betterness, and even bestness, then it is a necessary *a priori* truth that the act of teleological comparison of tendential necessities is better and is, indeed, the best function. If there is a valid and sound basis for stating that something is better than something else, then we are logically compelled to accept that basis. If, therefore, it is the case that the act of teleological comparison is necessary for betterness and that it necessarily yields betterness, *we must admit that the act, itself, of teleological comparison of tendential necessity is intrinsically, objectively better than all those things which are not the act of teleological comparison of tendential necessity.*

6.1.18 *Objections on Grounds of Inherent Relativity and a Reply*

One may pose a further objection. “You keep saying that we *must* narrow down a single *telos* in order to discover what is better and best, but what if such terms as “better” and “best” are *inherently* relative terms in that they can only be used with reference to arbitrarily chosen *teloi*?” First, let us see what such thinking would entail. If such terms as “better” and “best” are *inherently* relative terms in that they can only be used with reference to arbitrarily chosen *teloi*, then this, as I stated before, means that all value judgments would be equally sound. Once again, however, this thesis requires a rejection of the analysis in Chap. 3 on negative and positive justification. Since I have conclusively shown that there are, indeed, certain ethical judgments that are invalid (e.g.: “I ought to beat and rob person X because I want money”), and since Gewirth and Beyleveld have, indeed, shown that certain determinate ethical judgments are sound (e.g.: “I ought to help those who cannot help themselves if I am in a position to do so”), then it follows that there are, indeed, sound “ought”- and “must”-judgments. Furthermore, as we have seen, “ought”- and “must”-judgments are covert value judgments; such judgments entail judgments of goodness or value and, as Gewirth and Beyleveld emphasize, it does not matter if such judgments of goodness are prudential ones, instrumental ones, or moral ones. Indeed, as Smith and others also suggest, one cannot even make a prudential value judgment without making a judgment that one’s own prudentiality is objectively valuable. If it is not objectively valuable, then the whole point of prudentiality and especially egoism is questionable. Indeed, if there is no objective value, then even egoism can have no *point*, as it were. In fact, even nihilism can have no *point*; there can be no reason to believe that nihilism is true. Furthermore, due to much previous analysis, our objector cannot even appeal to the argument that values are merely desires and cannot therefore be true or false. The arguments of Lillehammer, Hampton, Smith, Foot, and others have shown that value claims cannot be mere claims of desire, or as such philosophers as Peter Railton argue, interests which take the form of second-order desires. Desiderative internalism, it turns out, is completely unsound and even somewhat paradoxical. There also exist other compelling reasons why value claims must be objective. In the first instance,

one can believe erroneously that values are a class of desires and still plausibly hold that there is such an entity as objective value. Hare, as we have already seen, conceptualizes values in just this way, but he uses the idea of normative validity to demonstrate that value claims have the capacity to contradict each other and thus lose their objective validity. Hare, of course, was wrong to think of values merely in terms of desire, but the fact that his theory nonetheless remains plausible in this respect is telling. Railton also conceptualizes value in terms of desire, although his approach is perhaps more sophisticated. Although for him, value is a kind of second-order desire, values also have the not inconsequential attribute of entailing certain attitudes. As he writes:

“If taking an attitude or embracing a desire involves certain beliefs, it also involves an assumption that these beliefs are not merely false, that one’s outlook is not largely a matter of, or psychologically dependent upon, error or ignorance. What is it about the possibility of error or ignorance that creates a potential threat to one’s outlook itself? The beginning of an answer may be this. We call upon our basic goals, as distinct from our mere (though perhaps insistent) desires, to explain to ourselves and to others the *worthiness* and *point* of our choices—indeed, of our lives.”¹³

Habermas’s discursive approach to ethics, furthermore, emphasizes the necessarily objectivizing nature of discourse. His entire theory centers on the phenomenon of objective value claims on the part of agents. When an agreement, any agreement, between two rational subjects is made, he argues, “The agreement is regarded as valid not merely ‘for us’... but as objectively valid, valid for all rational subjects.”¹⁴ In other words, if agent A and agent B agree that *x* is the case, then they implicitly expect that all agents, given relevantly similar scenarios and expectations, would agree that *x* is the case. The crucial point that Habermas wishes to make here is that, without claims to objective value, any kind of agreement concerning principles of value would be impossible and incoherent. Value agreement happens, however, all the time. Indeed, on Habermas’s account, the objectivization of value makes the very holding of a norm within the mind of a singular subject possible. If a subject holds a norm and acts on it, indeed if she does anything at all, she must accept that, if she engaged in discourse with other subjects in the Ideal Speech Situation, they would agree to hold her norm as well.

The objectivization of value, on Habermas’s account, makes the holding of a norm in the mind of a singular subject possible in yet another way. Since norms can only arise within the context of the lifeworld, and since the lifeworld in which the subject grows from infancy to adulthood is imbued with already established values and value systems, then the subject can navigate her or his own normative beliefs only via the proposed norms of other subjects. While the relativist would cite this as evidence that value claims lack objectivity, Habermas uses these facts in precisely the opposite manner. Contrary to what relativists might claim, that subjects develop their moral

¹³ Peter Albert Railton, “Facts and Values,” *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Ed. Peter Albert Railton. (Cambridge: Cambridge University Press, 2003): 53.

¹⁴ Habermas, *MCCA*, 292.

consciousness within the context of the lifeworld shows that the content of these norms is, at least to some extent, held as objectively valid on the part of the subjects themselves because it is integrally constitutive of their communicative interactions. In order to interact, hold a norm oneself, and convince others to hold the same norm (which makes possible the holding of norms on the part of other subjects as well), the dynamic discursive content that constitutes the cultural milieu must be objectivizing.

Existentialism, while mostly implausible, lends us another way of understanding the inherent objectivization of value on the part of agents. The existentialist notion that one's acting in the world changes the world in some way and thus forces the agent to accept some claim of objective permissibility stands here quite prominently. Any idea of objective value (the commensurability of *teloi*) is akin to this. A similar idea is comprehended somewhat in Habermas's theory of communicative action. It just is the case that one makes claims to objectivity when one makes statements of value, partly because one thereby makes an implicit statement concerning that attitude which all others should take with regards to one's action. In the scope of temporal sequence, one's actions unalterably change the course of events. One thus necessarily accepts this change for all others and prescribes for all others the acceptance of such a change. There is no escaping such a judgment.

Our objector, however, does not even have to accept my claims up to this point in order to accept my reply, for, if there exist objectively true "ought"- and "must"-judgments, as I have shown, and if such judgments entail positive judgments of goodness or value and vice versa, then it follows that there exist objectively true judgments of goodness or value such as "I ought to do X rather than Y if and only if doing X is better than doing Y." Since we have already demonstrated that objectively true judgments of this sort exist, then we have perfect license to reject the subjectivist and relativist theses. This, in turn, is why we also have good reason to search for a sound *telos* for betterness and bestness. According to my account, if we are able to make such objectively true value judgments, it follows that there must exist such a *telos*, for, if I can make objectively true value judgments, then it follows that there must be *some* sound standard for value judgment which grounds this judgment or which, in other words, makes the objective truth of this value judgment possible. Of course, it could be the case that there are some things that are supremely equally good, but this seems question-begging, for we must still ask, "equally good in relation to or compared to what?" We see then, that, if I can make objectively true value judgments, it follows that there must be some *telos* of bestness. Thus, if there is some *telos* of bestness, then we have perfect license to attempt to discover it. If we can find such a *telos*, we are, furthermore, obliged to accept it. All of this utterly destroys the subjectivist and relativist theses.

6.1.19 Objection on Grounds of Failure to Prove that Value Is a Natural Property and a Reply

The objection may be offered that I have not, strictly speaking, proved that goodness, betterness, or value more generally, is a natural property. It may be

objected, in other words, that I have demonstrated merely that such concepts and the positions concerning such concepts are necessary for agents to hold or value and not, strictly speaking, that value is “out there” in reality. First, let me be clear about what I imply through the use of the term “natural:” I do not imply, through the use of this term, physical existence as such. My use of the term is broader than this. Thus, when I refer to value as a natural property, I imply that value is a property of some entity in reality, whether that be in physical reality or non-physical reality. As we have seen, the concept of goodness comparison, or the comparing of goods relative to a single *telos*, is necessary for the concept of betterness. This concept of betterness, therefore, is instantiated by the concept of goodness comparison. There are clear grounds then, we must see, for the claim that betterness is a property of this activity of goodness comparison. Will our objector go on to deny that the empirically observed phenomenon of goodness comparison does not exist? It is only through such a denial, we must see, that our objector could refute the claim that betterness is a property of the world or, in other words, of reality.

The objection to the status of goodness and value more generally as natural properties is even easier to refute for, as we have seen, goodness is tendential necessity of function or, in somewhat simplistic terms, a property that an entity has which makes its fulfillment of some function more probable. How can our objector deny that tendency, necessity, and function so construed are properties of the world, that is, without denying certain other fundamental epistemological axioms concerning the world? I believe, therefore, that there is required the relief of a significant burden of proof in order to show that such entities as goodness and betterness are not properties of the world.

6.1.20 *Objection on Grounds of Dialectical Necessity and a Reply*

The objection might arise that mine is not, as it were, an assertoric theory. Indeed, one might charge, my theory is nothing but an elaboration of Gewirth’s or Habermas’s theory. Needless to say, we have not yet surveyed the entire theory, but enough has been elaborated at this point to enable me to respond to such a charge. Gewirth and Habermas, as well as Apel, set out to answer one form or another of a very specific question, namely “What must all agents admit on pain of contradicting that they or other agents possess some necessary property or are otherwise engaged in some necessary activity?” As such, their resulting respective methods are aimed at demonstrating those necessary statements or inferences that agents must make, *regardless of their actual truth*. Such a question is not that which I seek to answer and, as such, my method is wholly different. My approach can be seen as what might be called a transcendental reductionism: I attempt to discover the basic conceptual fundamentals of value and to trace all value judgment back to this ultimate source. The question I seek to answer, by contrast, is “What is the *meaning* of normative statements, and what concepts and conceptual frameworks are *necessarily* invoked by statements of value? What is the necessary conceptual framework, that is, that which underlies and

finds any and all conceptions of value?” Note that such a question is relevantly similar to the question “What is the *meaning* of physical statements, and what concepts and conceptual frameworks are *necessarily* invoked by statements of (let us say classical) physics?” Perhaps we shall answer that the ideas of position, velocity, and acceleration ground such a framework—perhaps even the very ideas of differentiation or integration—and this would require an equally thorough analysis. We must see, however, that this question, like the above question pertaining to value, is not a question of dialectical necessity but of assertoric necessity. We are not asking, in other words, regardless of the truth of the matter, what agents must admit with regards to the underlying assumptions of classical physics on pain of contradiction. We have brought ourselves, we must see, out of the confines of the dialectical approach and have truly ventured into the domain of the *post*-dialectical, possessing of vital insights and ideas central to its method but altogether independent of it. A partial indication of this can be seen in the assertoric method’s contribution to the greater applicability of value-theoretical concepts via its survival of the Single-Person Problem, and we shall see concrete proof of its survival in the next section. Its survival of this problem, as well as non-aretaic methods’ failure to survive this problem, also provides strong evidence to support the notion that traditional Western non-aretaic methods truly do need the greater explanatory power provided by virtue theory and a correlative theory of the good and the better if they are to prove capable of answering the deep normative questions of existence that lie beyond the shallow prompts “What are my rights?,” “What are my duties?,” and “What is just?”

6.1.21 *Objection on Grounds of the Intrinsicity-Finality Distinction and a Reply*

One might object that, because there is in fact a real distinction between intrinsic goods and final goods, it does not follow from my analysis that the intrinsic goods I have identified are also final goods. In her classic article, “Two Distinctions in Goodness,” Korsgaard argues in favor of making such a distinction. As she states:

There are, therefore, two distinctions in goodness. One is the distinction between things valued for their own sakes and things valued for the sake of something else – between ends and means, or final and instrumental goods. The other is the distinction between things which have their value in themselves and things which derive their value from some other source: intrinsically good things versus extrinsically good things. Intrinsic and instrumental good should not be treated as correlatives, because they belong to two different distinctions.¹⁵

Korsgaard’s overall argument, however, is not, strictly speaking, an argument *for* this position but, rather, an argument against the mere assumption on the part of value theorists that there exists no such distinction. The argument, therefore, takes more clearly the form of a proposal than a rigorous proof, and of course there is

¹⁵ Korsgaard, “Two Distinctions in Goodness,” *Philosophical Review*. 92, no. 2 (1983): 169.

nothing inherently problematic about this. Shelly Kagan, Wlodek Rabinowicz, and Toni Rønnow-Rasmussen, most notably, have taken up and accepted this proposal.¹⁶ I do not wish to venture too deeply into this debate, at least in this space, for reams of journal articles can be and have been written on this topic. The most I will do here, therefore, is attempt to demonstrate my theory's natural response to such a proposal.

I think it is clear from what I have demonstrated thus far that, if value is intrinsic, then it must also be final. Something cannot be good or otherwise have value without, I have been claiming throughout this work, appealing to and fulfilling some particular criteria which are themselves finally, non-derivatively good. This is indeed inherent in the concepts of teleologically comparative tendential necessity of function and reflexive intrinsicity. Because such concepts as betterness and bestness inherently instantiate a teleology and because they are intrinsic properties, therefore, then if they are the only entities which make any betterness or bestness at all possible, then it follows that anything that is intrinsically valuable is so because it makes the instantiation of something that is finally valuable more probable. I am once again, therefore, on the side of Zimmerman's position in repudiating this distinction.¹⁷ Based on what I have presented thus far, as well as on what I am to present, I cannot see any basis whatsoever for such a distinction and, thus, for such an objection.

6.1.22 Final Considerations: The Inadequacy of Summation and Aggregative Models of Value

Before moving on to an analysis of the implications of my conclusions regarding teleologically comparative tendential necessity of function and the act of teleological comparison of goods, it is worth making the final note that I do not offer, in my theory, any particular formula for the determination of the value of an entity. This is not because it is not possible to do so; indeed, it is certainly possible, in principle, to express my model of value attainment as a probabilistic one. I believe that such a formula, for reasons that will be seen in the next section regarding the relevance of experiential rationality, would be of very limited use. In any case, however, it is worth emphasizing that, if formulated, such a formula would emphatically not be based upon a method of quantitative aggregation or summing. The assumption that such a method is valid, however, has gone virtually unchecked among value

¹⁶ Shelly Kagan, "Rethinking Intrinsic Value," *Journal of Ethics*, 2, no. 4 (1998): 277–297; W. Rabinowicz and T. Rønnow-Rasmussen, "Tropic of Value," *Philosophy and Phenomenological Research*, 66 (2003): 389–403; "A Distinction in Value: Intrinsic and for its own Sake," *Recent Work on Intrinsic Value*. Ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman. (Dordrecht: Springer, 2005): 17–31.

¹⁷ Zimmerman, "Intrinsic Value and Individual Worth," *Recent Work on Intrinsic Value*. Ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman. (Dordrecht: Springer, 2005): 191–205.

theorists. Utilitarianism, of course, is the quintessential example of such a method, but classic examples are manifold in the modern and contemporary value-theoretical literature as well.¹⁸ Most of the models proposed in this literature base themselves upon one or another interpretation of expected-value theory, and this, in itself, constitutes a significant normative assumption. Interestingly, Neil Feit is one of the few value theorists who have eschewed purely aggregative models and have pursued, instead, mathematical or probabilistic modeling in which aggregation of value is not the central feature.¹⁹ His method of ranking values by ordered pairs, while still incorporating a summing method, at least acknowledges the existence of a non-aggregative manner in which higher and lower values can be ordered. Thus, the approach is promising, but, as do the others, it leaves itself open to the charge that the method of summing or aggregating value is merely assumed to be valid. This, however, is a very serious concern for, even in the field of statistics (the field in which such methods of modeling are primarily operative), there is an informal debate concerning whether or not it is legitimate merely to assign numerical values to qualitative variables and proceed with the usual tests.²⁰ Another approach which I believe to be promising, but which ultimately faces the same problem as Feit's model, is that of Daniel Doviak. His virtue calculus, while accounting for the notion of degree of goodness and badness, uses a method of value assignment that I believe to be highly problematic. He states:

The IVV+ for a given action is the sum of the values assigned to that action by the various virtue-indexed scales. Thus, if an action gets a .2 on the justice scale, a .4 on the benevolence scale, and 0s for each of the other virtues, then this action's IVV+ will equal $(.2 + .4) = .6$.

The IVV- for a given action is the sum of the values assigned to that action by the various vice-indexed scales. For example, suppose that the same action as the one mentioned above gets assigned $-.3$ on the cowardice scale, a $-.1$ on the infidelity scale and 0s for each of the other vices. Then this action's IVV- will equal $(-.3 + -.1) = -.4$. Since the action's net-IVV is the sum of its IVV+ and IVV-, the action's net-IVV will equal $.6 + (-.4) = .2$.²¹

His "virtue-indexed scales," however, fall prey to a problem I have previously argued is inherent in all pluralistic accounts, namely that of value incommensurability. His calculus, in other words, asks us to quantitatively compare differing

¹⁸ G. Arrhenius, "Superiority in Value," *Recent Work on Intrinsic Value*. Ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman. (Dordrecht: Springer, 2005): 291–304; Erik Carlson, "The Intrinsic Value of Non-Basic States of Affairs," *Philosophical Studies*. 85, no. 1 (1996): 95–107; R. M. Chisholm, "Intrinsic Value," *Recent Work on Intrinsic Value*. Ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman. (Dordrecht: Springer, 2005): 1–10; Sven Danielsson, "Harman's Equation and Non-Basic Intrinsic Value." *Uppsala Philosophical Studies*. 47 (1998): 11–19; Fred Feldman, "Basic Intrinsic Value." *Philosophical Studies*. 99, no. 3 (2000): 319–346.

¹⁹ Neil Feit, "The Structure of Higher Goods," *Southern Journal of Philosophy*. 39, no. 1 (2001): 47–57.

²⁰ Special thanks to Columbia statistician Anthony Donoghue for introducing me to this debate.

²¹ Daniel Doviak, "A New Form of Agent-Based Virtue Ethics." *Ethical Theory and Moral Practice*. 14 (2011): 268.

scales without providing an account of how such quantitative comparison *between* scales is possible. If, however, we choose to refer merely to *one* scale, thus upholding a monistic theory, we become capable of eschewing such an unnecessary theoretical complexity. This, however, is possible only in invoking the idea of *telos* and, perhaps, in therefore adapting Doviak's theory to a model of probabilistic comparison with respect to an ultimate *telos*, the *summum bonum*. Fortunately, Doviak seems to welcome such innovation.²²

As I believe my analysis has thus far shown, however, aggregative models are seriously problematic and most likely derive from a conception on the part of most of such theorists of value as some form of desire. Indeed, I believe this is evident from their writings. We have seen, however, that all essential meaning of goodness, betterness, and bestness is inherently devoid of any reference to desire or desirability, let alone intrinsic desirability. Thus, if this is the *prima facie* justification set forth in favor of aggregative models of value, then aggregative models most certainly fail. In any case, in lieu of an aggregative model of value, we have seen, at least thus far, that a model of value judgment and comparison cannot find its fundamental basis in mere aggregation. We have seen, rather, that some form of probability calculation is at work but that it is not one in which the value of some entity or state of affairs is dependent upon the probability value itself. Instead, the value of such an entity or state of affairs is derived from the *telos* itself in question and not, as seems to be the case in most aggregative models, upon the final tallying of values yielded by the model itself. In addition to desiderative conceptions of value, therefore, I believe that another reason why such aggregative models are so prolific is that such theorists have failed to consider the teleological nature of value judgment. As a result, the modern attempt to make value judgment intelligible without teleology or, otherwise, an appeal to the concept of *telos* seems to have taken the form of a numerical assignment to different value states and a derivative assumption that aggregation of these numerical assignments, in one way or another, is the only possible method of modeling value. As we progress to the next section, we shall see more clearly the inherent implausibility of such aggregative models.

6.2 The *Summum Bonum*

In this section, I demonstrate that, from all of my previous analyses, we can logically conclude that there is, indeed, a *summum bonum*: best rationality, wisdom, or alternatively "knowledge in a state of total freedom." I thoroughly define and characterize this *summum bonum* and conclude that, while Gewirth and Habermas's theories are consistent with the theory I set forth, my theory is multi-dimensional in that it does not leave out the important question of what goodness and value essentially are. As part of a special grammar, I argue, the language of goodness

²² *Ibid.*, 271.

and value just is the language of rationality. As such, four fundamental values, which I also exhaustively analyze and define, derive from the *summum bonum*: *Phronesis* or Best Logical Judgment, Best Conceptual Abstraction, Best Conceptual Synthesis, and Best Freedom or Liberation. While *phronesis* comprises the deductive and inductive components of rationality, the three latter components are experiential and phenomenological. As such, important virtues, high among them self-knowledge, derive from this *summum bonum*. While seemingly very abstract, I explain the immediate practical applications of this in an individual life and show concretely how to apply the theory.

6.2.1 From Value to Rationality: The Fundamental Link Elucidated

By opening up the world of betterness and bestness, we open up a whole slew of things that we can now compare to each other. What, exactly, can we conclude? The most obvious thing we can conclude is that anything that is necessary for the function of teleological comparison of tendential necessity is good and, more importantly, that, if function X is more tendentially necessary for the function of teleological comparison of tendential necessity than function Y is, that is that function X has a greater tendency to lead to or yield the function of teleological comparison of tendential necessity than function Y does, *then X is better than Y*. We will see exactly why this is relevant shortly. First, let us consider something else that is momentous: If logical rationality and experiential rationality are necessary and sufficient for each other as we discovered in Chap. 2, then, whatever is true of logical rationality is also true of experiential rationality. Furthermore, if logical rationality and experiential rationality are both necessary for the function of teleological comparison of tendential necessity, or even if, for the sake of argument, only one of them is necessary for the function of teleological comparison of tendential necessity, then it follows from our analysis that it is the case that at least one of the functions of rationality (logical or experiential) is not merely good for but best for teleological comparison of tendential necessity. If it is the case that at least one of these two functions of rationality is best for teleological comparison of tendential necessity, then, from bi-conditionality, it follows that it is the case that both functions of rationality, that is both the logical and experiential functions, are best for teleological comparison of tendential necessity. We must recognize, then, that we are not speaking of teleological comparison of tendential necessity in a conceptual vacuum; teleological comparison of tendential necessity is inextricably logically tied to rationality. What follows from this is equally momentous. If it is the case that teleological comparison of tendential necessity is the best function that exists, *then it is an analytic consequence that rationality is the best function that exists*. This is why, perhaps, it does not seem to make a whole lot of sense to make such statements as “Evaluation has no value or is valueless” or “Non-evaluation is

more valuable than evaluation.” If we analyzed these statements as well, we would surely find contradictions. Let us, now, return to the above statement: “The most obvious thing we can conclude is that anything that is necessary for the function of teleological comparison of tendential necessity is good and, more importantly, that, if function X is more tendentially necessary for the function of teleological comparison of tendential necessity than function Y is, that is that function X has a greater tendency to lead to or yield the function of teleological comparison of tendential necessity than function Y does, *then X is better than Y.*” We see, now, that we can justifiably replace the expression “teleological comparison of tendential necessity” in this statement with the word “rationality.” What does this, then, entail? Let us consider those things that are necessary for rationality, namely the two functions to which we have consistently been referring: existence and life. Existence, we have already shown, is intrinsically good, but what things that exist are better than mere existence? We might compare a rock and life. Life, we can justifiably say, is better than a rock, since life has greater tendential necessity to yield rationality than does a rock. This is something we can empirically verify or, if you prefer, falsify, since it can be demonstrated that something that is living is more likely to be rational than something that is a rock; living things are more likely than rocks to have rationality. It can justifiably be stated, thus, that, *ceteris paribus*, it is better not to kill than to kill and, subsequently, that, *ceteris paribus*, one must not kill. These sorts of claims, we can now see, can be shown to be necessarily true. One might object, saying, “Life is not actually necessary for rationality. One can, by the scientific definition of life, not have life and still have rationality since it is conceivable that artificial intelligences can be rational and not, strictly speaking, be living.” This, however, is beside the point. The conception of life of which I am speaking, however, is much more expansive and encompassing. Living things, as I am broadly defining them, are all those things which have interests. It is not meaningful, for instance, to state that a rock has interests. It is, however, meaningful to state that birds, tigers, and even perhaps ants have interests. Those living things have interests which have things which are valuable for them. All living things consistently undertake actions to continue their own existence. This, then, includes such entities as ants, flowers, and bacteria. One might object, however, that many entities that we call living, including humans, often undertake actions to discontinue their own existence. This, however, misses the point. Evolutionarily, all living things, including bacteria, have developed such that they either nourish themselves, release wastes, or undertake other actions *at some regular interval* which result in their continued existence. It is of no consequence to the argument, therefore, if, at some point in the entity’s life, it undertakes actions to kill itself or to let itself die. We realize, then, that the objector does not get around the requirement, for instance, of not killing. This, we must see, is because having interests is common to all living things, including things that have rationality or, we should say, agents; having interests, Hare might say, constitutes a criterion of relevant similarity. To voluntarily kill, maim, or injure something, thus, entails the judgment, *ceteris paribus*, “It is better to kill, maim, or injure something that has the interest not to be killed, maimed, or injured *and* it is better that my interest not to be killed, maimed, or

injured be unhindered and unmolested.” Although criteria of relevant similarities might be useful to prove my point, it is not actually necessary. As we have seen, we actually have the tools available to us to demonstrate that “It is better to kill, maim, or injure something that is living (than to refrain from doing so)” is false and, thus, unjustifiable. If life is better than non-life by virtue of its necessity for rationality, then anything which deprives living things of those virtues which are necessary or tendentially necessary for their continued survival and the good life characteristic of their respective species is worse, and not better, for life and, thus, worse, and not better, period. Why did I include the expression “the good life characteristic of their respective species?” I included it because not maiming or not injuring is not necessary for survival, at least in all cases, but we can make a strong case that not maiming and not injuring is more tendentially necessary for life than maiming and injuring. Let us demonstrate this. I used the expression above because it closely fits Foot’s virtue-theoretical paradigm and its focus on Aristotelian categoricals and the teleology of species. As we did with our examples, in previous pages, of “good for a thief” and “better for a thief,” we would simply have to discover the living thing in question’s essence, and then discover its *telos*, which we would discover by observing which of the organism’s functions tend to yield survival. After we have done this, we will be able to judge, at least with a high degree of accuracy, which activities would tend to hinder, instead of yield, survival. Certainly, most organisms that are maimed or injured are in a substantially or at least moderately weakened and vulnerable state which leaves the organism less able to provide for itself the things it needs for continued survival, such as protection from predators or the ability to undertake mutualistic behavior. We have seen, then, that, contrary to what some would like to claim, rationality’s status as the best existing function does not exclude non-human animals, or even grass and trees, from the realm of good action, although we will have to admit that, when the same interests of a non-human animal and an agent conflict, say for instance life, *ceteris paribus*, agent interests take priority. What this theory of the good has seemed to demonstrate is that all living things, and indeed all things that exist, deserve our respect. We must, therefore, have respect for nature not because it provides us with some direct benefit, as Kant and others would lead us to believe, but because it is good in itself.

One might, at this point, object on the grounds that this theory of the good is elitist, since it is not the case that everyone can be rational. If what is meant by this is that it is not the case that all agents can be rational, this is a contradiction. If, however, what is meant by this is that it is not the case that all humans can be rational, then this makes a little bit more sense. We might say, for instance, that a human in a persistent vegetative state is no longer an agent and is, thus, no longer rational. This, however, only has the consequence that this human being is a *worse* human being because it has been made *worse* by some tragic situation. As a worse human being, however, it is still a living being, and, as a living being, it deserves the same respect we would give to any living being. There are, of course, degrees of tendential necessity. This human being deserves more respect, for instance, than that which we must give to grass, bacteria, ants, and almost all other animals. Why is this? It is because this human being, although it is in a persistent vegetative state,

is comparatively more tendentially necessary for rationality than most other animals, although we might consider such beings as bonobos, octopi, and dolphins as, perhaps, comparatively more tendentially necessary for rationality than a human being in a persistent vegetative state. This would have to be resolved empirically. Living human beings, therefore, deserve our respect in the form of proper care, comfort, and dignity, and they deserve this respect because living human beings have the degree of intrinsic goodness that they have.

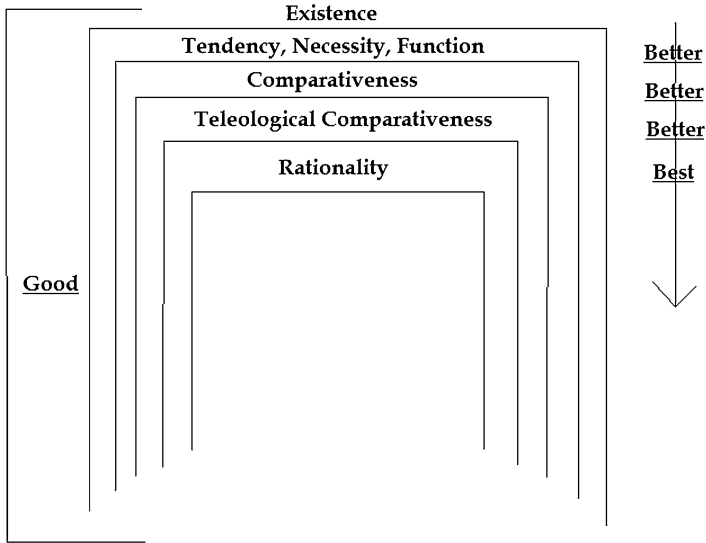
Our objector might be persistent. “Your theory is still elitist,” she might say, “since a logical consequence of the theory is that people who are more rational than others are better than others and, therefore, have more rights than others.” This kind of a comment demonstrates exactly how careful we must be with our logical inferences, since this is actually an invalid inference; none of what she stated follows from the theory. First off, it is not possible to be more or less rational than another person; an individual is either rational or it is not. In other words, individuals either exhibit the relationship,

$$\text{HasLogicalJudgment}(\Phi) \leftrightarrow (\text{HasConceptualSynthesis}(\Phi) \wedge \text{HasConceptualAbstraction}(\Phi) \wedge \text{HasFreedom}(\Phi))$$

or they do not. If, however, what is meant by this comment is that a person uses these faculties more efficiently than another person or that a person is more intelligent than another person, whichever most likely flawed criteria we are to use to define intelligence, then he is claiming something else entirely. He is making a right-claim concerning intelligence or better use of faculties. Gewirth and Beyleveld, however, have dealt quite adequately with such a claim. From Gewirth’s theory, and especially from Beyleveld’s further elucidation of the Argument from the Sufficiency of Agency and the Argument from Attitudinal Consistency,²³ we can see that agency, itself, is sufficient for the possession of negative and positive rights to freedom and well-being. Someone’s intelligence, just like someone’s race, gender, sexuality, socioeconomic status, or eye color, is an impossible basis for a right-claim. The reader can see this, however, without all of these fancy arguments. If the reader simply revisits Chap. 3, Sect. 3.2, on developing a method of justification, she will see that such criteria fail the test of constituting sufficiently justifying reasons. Thus, these charges of elitism are null.

To get a good sense of how the argument has developed thus far, I have provided a figure below which demonstrates the nested nature of the frameworks we have mentioned, their inherent status in what Foot has referred to as the “grammar of goodness,” and their derivative betterness. The usefulness of linguistic-empirical frameworks helps us see this development graphically.

²³ Beyleveld, *The Dialectical Necessity of Morality*, 95.



6.2.2 Justifying a *Summum Bonum*

Let us now explore the practical consequences of rationality's status as the *summum bonum*. Are we to conclude, for instance, that the best kind of agent life is a life revolving around *theoria*, as it is with Aristotle? Perhaps its status as the *summum bonum* entails wisdom in accordance with the goal of *ataraxia*, or peace of mind, as was the case with Epicurus. Epicurus, however, along with many classical and hellenistic virtue theorists, believed that the good life of a person was inextricably bound up in doing good unto others. Thus, we see a familiar pattern in classical and hellenistic virtue theory, namely that wisdom, justice, and good living, are necessary and sufficient for happiness or fulfillment. In the case of Epicurus, these latter things are necessary and sufficient, specifically, for pleasure. As he states:

It is impossible to live a pleasant life without living wisely and well and justly, and it is impossible to live wisely and well and justly without living pleasantly. Whenever any one of these is lacking, when, for instance, the person is not able to live wisely, though he lives well and justly, it is impossible for him to live a pleasant life.²⁴

What we might think of as the best or most fulfilling pleasure is, for Epicurus, true pleasure. Indeed, even John Stuart Mill distinguished himself most markedly from Bentham and his father, James Mill, by invoking "higher" and "lower" pleasures, indicative of a covert importation of a very non-hedonistic standard—

²⁴ Epicurus, *Principal Doctrines*, Ed. Robert Drew Hicks. E-Text: <http://ebooks.adelaide.edu.au/epicurus/doctrines/>. 2007.

that of rationality—into his theory,²⁵ and it was Aristotle who notably remarked that pain and wisdom often coincide but that wisdom is good nonetheless. The point, here, is that, in virtue theories, most conceptions of a person's *telos* include, as inherent in self-fulfillment and as necessary and sufficient for it, acting well and, more precisely, acting virtuously toward others. Seek that which is good for yourself, they imply, and you will do good for others. This, however, has the potential to be sorely misinterpreted, as it has been by the so-called objectivist, Ayn Rand. Let us not, however, fall into the swarm of Randian fallacies. What this truly means is that one's own good and others' good coincide. There are not, on the view of many virtue theories, brutal divides between one's individual good and the common good; these, in fact, are the same interests. We must, however, be more specific. Foot reminds us that "to speak of a good person is to speak of an individual not in respect of his body, or of faculties such as sight and memory, but as concerns his rational will."²⁶ What, however, is good concerning our rational will? We know that rationality is good and that, in fact, it is the best good, but there must be some way of judging between better and worse uses of rationality. One could be rational, for instance, while devising an evil plot to stage a military coup, but one might also be rational while doing such things as feeding a young child. We have already established that it is better to preserve and respect life, but how does such an act coincide with our rationality? We can begin to reveal this by, first, considering that the act of choosing well necessitates consistency in judgment. To discover why good choice necessitates consistent judgment, we must, first, consider the logic of choice, itself, within the context of value judgment.

When we make a choice, we have seen that we always make choices in the form of "X is better/less worse than Y; therefore, I ought to do X." Since choices are of this form, we must notice, again, how choices are capable of being justified: by appealing to their truth values. A choice will only be assertorically justifiable if my statement "X is better/less worse than Y, therefore, I ought to do X" is true. We can see that it is tautological to state that we are categorically obligated to make the better choice in any given situation. It is tautological because "I must make the better choice" entails "It is better to choose what it is better to choose" since "I must make choice X" entails "It is better to choose X" and "better choice" entails "what it is better to choose." If "I must make the better choice" is necessarily true, and if "to choose well" follows from "to make the better choice," then it follows that "I must make the better choice" entails "I must choose well" and "It is better to choose well." To choose well is to make the better choice; this is necessarily true. By now, however, we might notice that there is yet another necessary truth here, namely "I must choose to choose" or "It is better to choose to choose." Although tautological, these statements are actually quite revealing. What they reveal is that it is a necessary truth to state "Choosing is better than not choosing." This is a necessary truth because, if "I must choose the better choice" is necessarily true, and choosing

²⁵ John Stuart Mill, *Utilitarianism*, (Indianapolis: Hackett Publishing Co. 2002).

²⁶ Foot, *Natural Goodness*, 66.

is necessary for choosing the better choice, then “I must choose (in order to choose the better choice)” is necessarily true. What follows from this is that choosing to choose is necessary for making the better choice and that, thus, choosing to choose is a better choice than not choosing to choose or even choosing not to choose; “It is better not to choose than to choose” is unjustifiable. We can more easily see this if we consider the aforementioned tautology “It is better to choose what it is better to choose” and its contradictory “It is not the case that it is better to choose what it is better to choose.” Within the tautology are the statements “It is better to choose X” in the antecedent and “X is a better choice” in the consequent. If we consider that all of our choices necessitate this sort of judgment and that we must, thus, also admit the necessary truth, “Choice is a necessary good (is necessarily good for) for making the better choice,” we must, then, admit the following:

1. Choice is a necessary good for making the better choice.
2. We make all of our choices by judging what is better.
3. Premise 1 entails choice’s being better than non-choice for making the better choice.
4. Making the better choice is a necessary good for any and all good or better voluntary action.

Conclusion: Therefore, choosing to choose is better than choosing not to choose, and I must choose to choose, and it is unjustifiable to choose not to choose.

This seems like a lot of trite analysis, but it is quite important if we intend to be meticulous, careful, and clear. It is clear, from here, that we must now ask the question, “What constitutes a better choice, and how can we justify it as a better choice by demonstrating it to be a better choice?” The first clue to answering this question is by examining, yet again, choice itself, as well as the function of “must” in decision-making. “Must,” yet again, ultimately designates categorical obligation. What, again, do we mean by this? We mean that it is a deontic modal operator of necessity that functions specifically within practical judgments and which is over-riding and standard-setting; it sets a standard rather than merely referring to it. Since these judgments take the form of conditionals, as all statements of necessity do, we can rephrase the “must”-judgment as “It is necessary that X.” The next question then arises, “What is necessary for what?” Our “must”-judgment, “I must choose well” or “It is necessary to choose well,” obviously picks out choosing well as the necessary condition. But what is choosing well necessary for? In short, choosing well is necessary for making a better choice. But what is necessary for choosing well? In other words, what must we do, or what is it categorically obligatory that we do, in order to choose well? In order to choose well, it is necessary to do all those things which at least promote choice in the first place: we must have freedom, we must have conceptual synthesis, we must have conceptual abstraction, and we must have logical judgment. If all of these things are necessary for choice in general, then it follows that all of these things are, in some respect, necessary for choosing well. It seems, then, that in order to determine what

is necessary for choosing well and not simply for choosing in general, we need to determine what additional necessary conditions apply to choosing well. In order to choose well, it is evident, it is necessary to exercise those faculties necessary for choice in general but, additionally, to exercise them well. Just as we did with choice, we must determine the function of these faculties in order to determine what constitutes their functioning well. The function of choice, we saw, was to pick out a better or more valuable course of action. And so, we saw, it is an analytic truth that choosing well is a better choice. Let us examine the function of these related faculties one by one and determine what makes them function better.

6.2.3 *Logical Judgment*

The function of logical judgment is to judge using logic. Agents all use at least a minimal amount of logical judgment, but it is not the case that all agents use their logical judgment well. Some agents, for instance, make invalid inferences or inaccurate inductive judgments. But what makes this *bad* logical judgment? Let us see. If the function of logical judgment is to judge using logic, then the function of good logical judgment is to judge well using logic, but we must get more specific than this. Consider that a certain type of judgment is necessary for choosing well, namely the simple tautology “The better choice is to choose well.” In order to know, for instance, that it is better to choose well than to choose badly in the first place and, thus, in order for us to choose the better choice, we must justify the claim “The better choice is to choose well.” In other words, in order to choose well, we must first demonstrate the claim to be true via the use of the laws of logic; we must be consistent in our judgments, we must not contradict ourselves, etc. In other words, we are wholly justified in stating that, in order to choose well, we must judge *correctly*. If judging correctly is necessary for choosing well, then we are also wholly justified in stating that, in order to choose well, we must judge well and that *judging correctly constitutes judging well*. Since one’s choices derive from one’s judgment, and since good judgment concerning one’s choices consists in picking out the justifiable claim concerning one’s choices in accordance with the laws of logic, then the better choice is that choice which derives from the better judgment. Thus, we have explained how logical judgment can be good or bad, better or best. The virtues which one must exercise, then, for the best possible logical judgment one can attain are those things which are the most tendentially necessary for the development of the best judgment one can attain. These virtues will be discussed after we have discerned the best functioning of the other rational faculties. As we shall see, there are a lot of commonalities, and there is a lot of overlap between these faculties with regards to the virtues. An interesting feature to note, before we move on, is that, so far, virtue theory seems to be consistent with the logic of the relationship between these different faculties. Recall that, due to the bi-conditional relationship of these faculties to each other, whatever is true of one of the faculties is true of all of them. If the faculties of conceptual synthesis, conceptual abstraction,

and freedom are necessary and sufficient for logical judgment, then, if it is the case that there are some virtues which are best for judging well, or what we may now permissibly call *phronesis*, then it must be the case that those same virtues are best for good conceptual synthesis, good conceptual abstraction, and good freedom. If the antecedent of a bi-conditional contains a claim concerning the greatest tendential necessity of some entity, then the consequent must also contain a claim concerning the greatest tendential necessity of the entities to which tendential necessity can be applied. What we begin to see, then, is the early stages of a verification of the Aristotelian, Epicurean, and even Cynic and Stoic claims that the cultivation of virtue in oneself is necessary and sufficient for right or good action toward others.

Let us note one more salient feature of logical judgment before we move on. When we make the judgment “I must judge well,” which, as we should acknowledge, by now, is necessarily true, we do not mean simply “I must judge well one time” or “I must judge well when I feel like it.” Such judgments, remember, are categorical, and the implications are huge. “I must judge well,” rather, is necessarily true for every judgment and every choice we make. If we choose well, we must judge well, and in every choice we make, it is necessarily true that we categorically must choose well. Thus, “I must judge well” is true for every choice we make. This is significant, for this implies that we must always be striving toward the attainment and exercise of the virtues, and particularly of *phronesis*. In other words, we can never judge rightly that it is better that, for some specific instance, we fall into vice. Demanding? Yes. Onerous? No. As we progress in our analysis, it will become clearer why this is the case. Striving toward our ideal self might be challenging, difficult, and sometimes frustrating, but such striving, it can now more clearly be seen, is what makes human life worth living.

6.2.4 *Conceptual Synthesis*

Conceptual synthesis, we might recall, is necessary and sufficient for logical judgment. It is, furthermore, a constituent faculty that is essential to experiential rationality, that is, that aspect of rationality which comprises the totality of those faculties which make agent experience possible. Conceptual synthesis, then, along with the other faculties which comprise experiential rationality, is central to our experience. What, however, constitutes good conceptual synthesis or conceptually synthesizing well? We know, first, that *phronesis* is necessary for conceptually synthesizing well and vice versa, whatever it is. In order to synthesize concepts well, we must, then, exercise good judgment and vice versa; we must, in other words, make correct or accurate judgments concerning various states of affairs with which we are presented in our experience. Our conceptual synthesis, then, must be both efficient and diverse. What does this mean, and why does it constitute good conceptual synthesis or, even more precisely, the best conceptual synthesis? Efficiency in conceptual synthesis, generally speaking, is the capacity to construct

concepts from other concepts, and so forth, quickly and with the least effort. Diversity in conceptual synthesis, by contrast, is the capacity to construct new concepts, or what are new concepts to the individual in question, from various other concepts. Why do these capacities or functions constitute good conceptual synthesis? This is because, even at some minimal level, efficiency in conceptual synthesis is necessary, but not sufficient, for *phronesis*. This is because judging correctly or accurately necessitates some level of quickness, as well as low effort. To some, this might seem strange; would not judging correctly or accurately necessitate the greatest possible effort? Strictly speaking, no. The greatest possible effort in conceptual synthesis is not a necessary condition of judging correctly, and it would be difficult to demonstrate that it is even a sufficient condition. Why, then, is the least possible effort necessary for judging correctly and, thus, for good and best conceptual synthesis? The least possible effort in conceptual synthesis is necessary, quite simply, because it is necessary for efficiency in conceptual synthesis. When one starts out in one's conceptual synthesis, one has to expend a comparatively greater amount of effort to synthesize and construct the simplest concepts. As one typically progresses, however, the same sorts of syntheses require comparatively less effort. The point is that such progression is necessary for the making of correct or accurate judgments because the ability to judge in this way is dependent on the ability, in some respect, to analyze the constituent criteria of any situation and to incorporate the results of such analysis into one's judgments, whatever those judgments are. The more efficiently I synthesize the constituent criteria, be it via syllogism or other processes, and the less effort is required for me to efficiently synthesize such criteria, the more correct and more accurate judgments I am capable of making (in the qualitative, and not quantitative, sense of "more"). When I state that best conceptual synthesis is the most *efficient* conceptual synthesis, I am not, for instance, devaluing people with learning disabilities. Remember, everyone has an individual *telos*. "Most efficient," therefore, means "most efficient with regards to *telos* X." For instance, it is well known that people learn in different ways. Thus the ways in which the individual synthesizes concepts might be literally slower than other individuals, but this manner of synthesis might be *best* for this individual's best conceptual abstraction, logical judgment, and freedom, given the potentialities derived from the individual's essence. After all, the *kind* of concepts synthesized—perhaps even extremely abstract or unwieldy within an individual's mental processes—matter. Different individuals begin in different places conceptually. What matters is that these concepts are in accordance with one's *telos*.

This gets us to our other necessary condition of good and best conceptual synthesis: diversity in conceptual synthesis. I have been careful, here, not to say "complexity in conceptual synthesis," and we shall see why, although diversity of conceptual synthesis, I think, is probably sufficient for such complexity. In any case, it is vital for agents to have diversity in the concepts that they have at their disposal. This is because any correct or accurate judgment is going to require different, that is non-identical, concepts, even if these different concepts are those of "a" and identity, respectively, in the statement " $a = a$." Furthermore, the making of any correct or

accurate judgments concerning one's own or others' goods requires the capacity to synthesize sometimes wildly different concepts and, again, the capacity to do this efficiently. We are justified, then, in stating that good conceptual synthesis consists in efficient and diverse conceptual synthesis. The best conceptual synthesis, then, consists in the greatest possible efficiency and diversity in conceptual synthesis. We begin to see, now, how the concept of experiential rationality and, thus, *good experience* becomes salient; diversity of experience is necessary for good conceptual synthesis, and a supreme diversity of experience is necessary for best conceptual synthesis. The theory, however, naturally discounts all harmful experience; the reason for this will become clear shortly.

6.2.5 *Conceptual Abstraction*

Conceptual abstraction and conceptual synthesis, predictably, have much overlap. Conceptual abstraction, needless to say, is necessary for conceptual synthesis. In order to do any synthesis or analysis, we must, in a very real way, imaginatively and creatively “detach modes of thought from raw sensory and perceptual input,” as Hauser's theory of humaniqueness states. Anything that is present for judgment in our experience, therefore, is subject to such abstraction and, indeed, such abstraction is necessary for any judgment we make. The most salient example of such abstraction, as Hauser's theory highlights, is the ability to apply one rule or set of rules that was learned in one situation and apply it to different, although relevantly similar, situations. The phenomenon of memory is, obviously, necessary for such abstraction. Such abstraction, in fittingly more abstract terms, constitutes the capacity to deconstruct concepts into their constituent parts and reapply these constituent parts in sometimes wildly different ways. This, of course, is vital for *phronesis*. How, for instance, can anyone be certain that one is making the correct or accurate judgment about something unless she has deconstructed and reapplied concepts in this way? This deconstruction and reapplication of concepts is necessary for *phronesis* because such deconstruction and reapplication enables us to more intimately understand these concepts and their logical, emotive, psychological, or phenomenological relationships and interrelationships. Such intimate understanding, it is clear, thus enables us to make more correct and more accurate judgments (again, in the non-quantitative sense) concerning the phenomena with which we are presented in our experience. Let me now make clearer a point which I made at the end of our discussion on good conceptual synthesis, namely that “the theory... naturally discounts all harmful experience” with regards to the development of good conceptual synthesis. This rule also applies to conceptual abstraction. Harmful experience is necessarily that which hinders the good exercise of conceptual synthesis or any of the other faculties, especially *phronesis*. Thus, we see that all harmful experience is naturally taken out of our consideration. For instance, it is possible to abstract too much. We can abstract so much that we lose all contact with others and simply stay tucked away with our thoughts and conjurings. This constitutes an exercise of bad judgment, and I will

explain why this is the case in due course. It is best for us to abstract, and to abstract as much as possible, but only insofar as it does not lead us to make bad judgments. Thus, when we say, for instance, that the best conceptual synthesis is the most efficient and most diverse, we are really saying that that conceptual synthesis is best which is the most efficient and most diverse with respect to *phronesis*. If the way in which we are exercising conceptual synthesis is bad for the exercise of good judgment, then the good exercise of both become hindered. This is the same for conceptual abstraction. We begin to see, also, the ideas of excess and deficiency that are central to Aristotelian virtue theory; it is always best for us to maintain an equilibrium or balance.

6.2.6 Freedom

This leads us to the phenomenon of freedom inherent in rationality. Bandura has given us a very expanded understanding of that in which freedom consists for agents. As we analyzed more in-depth in Chap. 2, agency is divided into various levels. We, first, have the base and behavioral levels which are the least voluntary in that their function involves the least choice. These are the levels which yield the semi-autonomous swatting of a fly with one's hand, the semi-autonomous spitting out of something bad-tasting, and other such semi-autonomous reactions. It does not follow from their semi-autonomous nature, however, that these levels are not subject to regulation or modulation. The reflective level, we might recall, is where the application of self-standards takes place; it is inherently self-evaluative. This level, in addition, has a meta-level motivational structure in that it interacts with the base and behavioral levels to regulate and modulate behavior; we might recall, for instance, CAPS theory and the concept of goal-dependent automaticity. Depending on the specific self-standard present in this level, a person might choose to reflect upon and develop a greater awareness of one's swatting flies or one's spitting out bad-tasting things in the future and, on the basis of the specific self-standard, develop the desire to work on not acting so semi-autonomously in such situations. This leads us to the next level, that of the self-referential level or, alternatively, the self-system. This level, we might recall, is constituted by self-conception and the motivations arising from a given self-conception. The level also has two interrelated subsystems. The self-evaluative system is the subsystem which interacts with the reflective level to actually *form* self-standards. Thus, while the reflective level merely acquires, maintains, and applies self-standards, the self-evaluative subsystem, and thus the self-system, works deliberately to shape and re-shape self-standards and self-evaluations. The other subsystem, we might recall, is the self-efficacy system, which is responsible for self-efficacy judgments, or judgments concerning one's capacity to do or accomplish certain actions or goals. That is why the self-system is self-referential: because all of the deliberation which takes place on this level exhibits self-awareness and, particularly, awareness of oneself as a moral agent. Rottschaefer emphasizes that the self conceived of in the self-system is the potential self contrasted with, or we might even say compared to, the ideal self constituted by the self-standards. This highest level of agency, I will demonstrate, is the best one and, thus,

constitutes the necessary condition for good freedom and the best freedom. Although those levels of agency below the self-system are necessary for the self-system, it is the self-system, in its totality, that is necessary for *phronesis*. Without the self-reflection inherent in the reflective level and the self-awareness inherent in the self-referential level or self-system, one would be far more likely to make incorrect or inaccurate judgments, which would constitute a failure to exercise *phronesis*. To make correct or accurate judgments concerning one's experiences, we must have and exercise those faculties of self-reflection, self-evaluation, and self-awareness inherent in the higher levels of agency; these phenomena are necessary for judging well. This is because self-reflection and the concurrent assessment of self-standards is constitutive of evaluative deliberation; the reflective level, itself, makes evaluation possible in the first place. The self-system is also necessary and is actually comparatively more tendentially necessary for *phronesis* than is the reflective level, since the self-awareness inherent in this level is responsible for setting standards in the first place, sweeping them away if perceived as inadequate or flawed, and replacing them with new standards. This, we must admit, is more tendentially necessary for *phronesis* and, indeed, is most tendentially necessary for it. The best freedom, we can now see, is that which is characterized by an exercise of the self-system that is necessary for the best exercise of *phronesis* or, in other words, the best exercise of good and best judgment. This, then, is the best freedom. Bandura's own characterization of freedom substantiates this. As he states:

A strong sense of personal agency requires the development of competencies, self-percepts of efficacy, and self-regulatory capabilities for exercising self-directedness. . . . These personal resources expand one's freedom of action and enable people to serve as causal contributors to their own life course by selecting, influencing, and constructing their own circumstances.²⁷

Frequent activation and exercise of the higher levels of agency, it can be seen, are necessary for such self-directedness. They are, in other words, necessary for being free of automatic response patterns and, more generally, of non-cognitive influences over one's behavior.

From this analysis, we can see then that there are different levels of freedom, each differing in their goodness from level to level. One form of freedom consists in merely doing what one wants or desires. Merely doing what one wants or desires, however, constitutes a superficial form of freedom. On this level of freedom, there is not much self-reflection or self-awareness; one remains constrained, as it were, by impulses that are not fully cognitive. Above this level lies the form of freedom that consists in doing what one believes to be valuable. On this level, there is a greater level of freedom because there is, among other activities, a greater level of self-reflection and self-awareness on the part of the agent because she or he has to consciously form beliefs concerning what is valuable. The best form of freedom that we have explored, however, is that which one can rightly call liberation. This freedom consists in doing what one values, with one's values being in accordance with the manifestation of best rationality according to one's individual essence and

²⁷ Bandura, *Social Foundations*, 38.

telos. As I proceed in my analysis, I will return to the subject of individual essence and *telos* and will examine how the faculties of best rationality play a role in constituting these. This concept of Liberation, in other words, will be more than satisfactorily elucidated. For now, let us follow some other pertinent leads.

6.2.7 *Wisdom and the Unity of the Virtues*

Throughout these analyses of the constituent concepts of rationality, we begin to see a theme, namely that *phronesis* is an extremely important, and one might even say fundamental, virtue. If one does not have *phronesis*, it will be impossible or nearly impossible to act virtuously. This, then, suggests the soundness of McDowell's Socratic thesis on the unity of the virtues as knowledge.²⁸ Although I would disagree with the thesis that it is knowledge, *per se*, that constitutes the unity of the virtues, it seems as though we can at the very least say that *phronesis* is a *united* virtue in that it encompasses all other virtues: all those virtues which are necessary and sufficient for a person's excellence are necessary and sufficient for *phronesis*. One can see a similar theme in Smith's work as well. His proposed linkage between the possession of good normative reasons and the possession of good motivating reasons under conditions of "full" rationality suggests my thesis. I believe, in fact, that the logical consequence of his analysis in *The Moral Problem* is that the greatest individual freedom corresponds to the greatest degree of correct logical judgment. Linda Zagzebski also suggests a similar role for *phronesis*. She first makes the very important distinction between skills and virtues. When applied to the intellect, such a distinction is vital, for while it is true that one may possess many intellectual skills, one may lack what I believe truly makes one intelligent: intellectual virtue.²⁹ Such virtue includes, as Zagzebski mentions, intellectual honesty and integrity, open-mindedness, willingness to listen to others' ideas and have one's own disproved, and others. The distinction between intellectual virtue and intellectual skill, she argues, partially explains why it is sensible to state that one might be both intellectually skilled and a morally bad person. Zagzebski critiques and then reinterprets Aristotle in going on to state, in fact, that it is impossible by contrast for one to be intellectually virtuous and morally unvirtuous or morally virtuous and intellectually unvirtuous. This is because the virtue of *phronesis* unites this distinction. My argument corroborates this as well. Zagzebski goes on to argue that *phronesis* in fact unites all virtues and stands atop them as a measurement of how virtuous a person actually is, in total. As she states:

If we are to avoid an excessive and unworkable fragmentation of value, there needs to be some virtue that permits a person to sift through all the salient features of the situation—

²⁸ John McDowell, "Virtue and Reason," *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. (Oxford: Oxford University Press, 141–163. 1997); *Mind, Value, & Reality*.

²⁹ I am essentially providing a normative reason for the re-definition of "intelligence."

that is, all those features that are pertinent to *any* of the virtues—and to make a judgment that is not simply the judgment of a person qua courageous, qua generous, or qua humble but is the judgment of a *virtuous* person. Therefore, the ability to mediate between and among the individual moral virtues must itself be a virtue. *Phronesis* is defined in part as the virtue that has this function.³⁰

I would go perhaps even further with this analysis. To attain *phronesis*, I contend, is to attain wisdom, and not just practical wisdom. Indeed, another way of stating that best rationality is the *summum bonum* is to state that wisdom is the *summum bonum*, for wisdom is, in fact, the best possible state of rationality. Through Zagzebski's distinction, we can clearly see that wisdom, characterized by virtue, is far more valuable even than mere knowledge alone, which can be and often is characterized by mere skill and which can thus be used for evil or base purposes. By contrast, there is no possible evil or base use of wisdom. This will become clear as the analysis progresses.

6.2.8 *Objections and Reply*

A major objection to my account thus far might arise. "It is not true," one might argue, "that best freedom is necessary for best or correct logical judgment; only some level of freedom is necessary." Another related objection might also arise, namely that correct logical judgment is indeterminate and, as a faculty regulating propositional attitudes, lacks teleological content. This, one might argue, makes the account vacuous, for correct logical judgment could concern anything, from seemingly the most trivial to the most important topics, propositions, or experiences. Because these sorts of misinterpretations are possible, I shall clarify a few ideas. "Correct logical judgment" is indeed meant to be taken in a general sense. According to what follows from the best exercise of rationality, namely best logical judgment or, as I have shown, correct logical judgment, it is best for rational beings to act in such a way that enables them to make correct judgments, or in other words acquire knowledge, about the world. As it stands, this account is far too general, for the question can still be asked "What *kinds* of things should I make correct judgments or acquire knowledge about?" The account, however, is meant to be general, and for a very important reason. As it stands, we are talking about the faculties of rationality as universals and not as applied to particular rational beings. This is helpful because it elucidates the general *telos* belonging to a *class* of things, namely rational beings. To give this account more direct content, as I will explain in other examples later as well, we would however need to apply this *telos* to a particular rational being with a particular essence. Now, given the question "What *kinds* of things should I make correct judgments or acquire knowledge about?" I can answer, "Examine your essence."

³⁰ Linda Trinkaus Zagzebski, *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. (Cambridge: Cambridge University Press. 1996): 222.

In the first case, this is for the purpose of discovering what it is possible for you to make correct judgments about. As it turns out, these possibilities are also myriad. Therefore, we must look to the other faculties for some assistance. We already know that best freedom is necessary for correct logical judgment, but this has been questioned in the first objection. One can correctly judge facts and states of affairs, the objector seems to hint, with only a minimal amount of freedom, for instance, at the reflective level. In fact, the objector might state, if I merely focus my attention toward learning as much as I can through rote memorization, I can easily attain the level of freedom I require to make correct judgments. Let us revisit, however, the fact that the *telos* of rationality, best rationality, mandates that I correctly judge both cross-temporally (all the time) and cross-situationally (in every way conceivable); we must remember that the claim is *general* and thus applies to all times and circumstances. This means that my goal in correct logical judgment is literally to judge correctly all the time and with regards to every conceivable subject matter, including all of the experiences and situations we face. This is indeed impossible for mere mortals to do, but it may not be impossible for some highly advanced rational beings of which we could probably have little comprehension. In any case, we must then come as close as we can to this goal. Let us see the implications of this. What is included in every conceivable subject matter? Among them includes a very salient one, namely behavioral standards. In other words, most salient in the search for correct judgment or knowledge is that of judging correctly how to behave. On a very minimal level, this includes the analytic judgment that it is evil to hinder the cultivation of best logical judgment (and this subsequently applies to other rational beings) and that it is good to aid in the cultivation of logical judgment (and this subsequently applies to other rational beings as well). There are, however, certain behaviors that one must accept as correlates to these judgments if one is really to judge correctly all the time and across situations. In order to judge correctly that these minimal judgments are correct, one must continue to judge this to be the case from one situation to the next and apply this knowledge to one's behavior in different situations. This is true in general, regardless of one's individual essence. Accordingly, a high degree of behavioral adjustment and personal standard-setting is necessary to achieve and maintain this behavioral commitment to helping others achieve knowledge and not hindering others in achieving knowledge. This level of behavioral commitment requires a high degree of self-analysis and thus a high degree of freedom such that one is able to judge and modulate one's own self-efficacy judgments and self-standards. This is the Bandurian conclusion made normative. We see, then, that, even on this minimal account, it is necessary for one to attain best freedom as I have characterized it. The minimal account, however, is just that: minimal. This is not the entire story. Recall that I stated that, in order to narrow down exactly which correct judgments one must make, it is necessary that one ask the question, "What *kinds* of things should I make correct judgments or acquire knowledge about?" We may recall that, for whatever essence an individual has, it is a necessary condition of judging correctly cross-situationally and cross-temporally that one attain best freedom. This is precisely because, for whatever one is judging, whether in the practical or theoretical realm, knowing one's own essence, or in other words knowing oneself, is necessary

for judging it correctly. In other words, while it is not necessary for judging *something* correctly that one judge it while in the state of best freedom or liberation, it indeed *is* necessary for judging *everything* correctly that one attain such liberation. Let me reveal a partial reason why this is the case. Consider that *everything* includes you. In order to judge correctly in the sense that I am stating it and in the sense in which it logically derives from best rationality, you must know how you think, how you perceive, how you analyze, how your memories affect your present judgment, how your ideas or evidence might be selective or biased, what states of mind are possible for you, and numerous other facts about yourself, empirical, phenomenological, mental, physical, and spiritual. If any of *these* judgments are inaccurate or incorrect, then knowledge of *everything* (or at least everything that otherwise would be available to you) is logically impossible for you. Knowledge of *everything* is a very tall order, but we must pursue it as far as we can. To make any progress in this endeavor, however, it is necessary for us to cultivate the virtues, and the only way we can do so is by first looking inward. It is through this elucidation, then, that the reader might see the connection to best freedom, best conceptual synthesis, and best conceptual abstraction. I want to focus specifically on freedom. This process of self-discovery via correct logical judgment coincides with a process of liberation, and vice versa. The question, however, remains: “What *kinds* of things should I make correct judgments or acquire knowledge about?” Now that we have cleared up some possible misinterpretations, the answer should seem obvious. Once again, we must look to our individual essences for the answer. Although it is implicitly part of every rational being’s *telos* to attain all knowledge that it is possible for her or him to acquire, a more specific answer is revealed by considering that, as the agent seeks correct logical judgment or knowledge via self-discovery and liberation, he realizes that the process of his slowly coming to know himself in a total sense has yielded a more specialized set of knowledge. This is due to the merely practical fact that it is only possible for us mere rational mortals to attain specialized knowledge about one, two, or several topics, situations, or states of affairs; we cannot have specialized, exhaustive knowledge about everything. In short, therefore, the answer to the question is that the knowledge an agent should attain is that which an agent attains in the process of gaining self-knowledge. If all of this is true, then it also follows likewise from this entire account that, as a general characteristic applying to all rational beings, correct judgment would include not only correct judgment about one’s essence but also the knowledge that one must not hinder, and indeed one must help, other rational beings in attaining liberation and self-knowledge. Of course, our judgments do not pertain to rational beings exclusively; there is a whole world of sentient and non-sentient beings which deserve our respect and consideration as well.

6.2.9 *The Question of Attainability*

We might stop to note now that, in all of the aforementioned faculties of rationality, the best functioning of these faculties is almost impossible for anyone, let alone

everyone, to attain. Everyone will, for instance, at some point, make a faulty judgment or do something harmful to himself or others. What, then, is the point? We must remember, however, that *teloi* are very often not attained by everything which has a *telos*. It is not the case that all zebras will run as fast as they should be able to run in order to escape lions or that a planet that has the conditions suitable for life will ever develop life at all. Our *telos* as rational beings, however, seems much more out of reach. How could I ever conceptually synthesize the best, conceptually abstract the best, have the best freedom, or exercise *phronesis* all the time? How, then, is it possible to exercise rationality the best? The fact is that it might, indeed, be impossible, at least for humans. We do not, for instance, know if there are extraterrestrials or other rational beings which exist which can achieve this *telos*. Alas, humans are imperfect, but we do not claim to be perfect. Who is to say that a poor mother in Somalia, Afghanistan, or the Bronx who lives to take care of her children, or even the oldest child whose mother has died who must now take care of his younger brothers or sisters, cannot, despite their tragic circumstances, advance toward their *teloi*? Certainly, from a Maslovian perspective, it is less probable that they will gain an education or attain any of those other goods that will enable them to exercise their reason to their best individual capacities, including their best exercise of conceptual synthesis, conceptual abstraction, *phronesis*, and best freedom or liberation, but they can at least *try* to judge the best that their circumstances allow and, perhaps, even exercise these other capacities the best they possibly can. Due to their severely undermined autonomy, they have severely limited choices concerning what their judgments can concern. It is clear, here, that their judgments will tend to revolve around their greatest tendential necessity for the exercise of rationality—that of life and survival. This raises an excellent question, namely “What is the agent *telos*?” Certainly we have seen that this *telos* is to exercise one’s rationality well, but such an exercise of rationality can take many forms. Which, we might ask, is the best form? Should we propose, as Aristotle does, that *theoria* is the best form and that, therefore, everyone should be a scientist of some sort? Perhaps we should, like Epicurus, take the position that we should simply work toward being self-sufficient, independent, surrounded by our friends, and mentally at peace? This, however, would be to over-simplify Epicureanism. Perhaps we should take a page from the poet and philosopher, Lucretius, who was himself an Epicurean. Lucretius addresses our concern of what *telos* among rational *teloi* is the best quite fastidiously in his *De rerum natura*. With regards to the issue of what kind of knowledge a person should pursue, Lenz explains:

What kind of knowledge? Knowledge of everything: of the way the world works, and in particular, knowledge of the hidden causes of all things. Lucretius’s poem *De Rerum Natura* explains physics (the workings of phusis, nature), psychology, epistemology, religion, myth, meteorology, and more. All this has implications for the human soul. Ancient thought did not separate science and humanities.³¹

³¹ Lenz, “How Epicurean Science Saves Humanity in Lucretius,” 89.

Lenz's point is instructive for many reasons. First, his critique, in general, is that of the modern distinction between science and the humanities. Yes, there is, perhaps, good reason to separate them in a taxonomical sense, but the point is that they are continuously overlapping and interweaving disciplines, and they are this way because they are part of the fundamentally overlapping, interweaving, and complexly unified phenomenon of rational experience. He continues:

Nothing is exempt from reason. The poet does not claim privileged access to something only he knows. He must persuade us. All human beings can have access to ultimate knowledge, and its rewards. While the goal is a high one that some might find to be elitist (on the lines of the well-known criticism of Plato)—having attained the highest knowledge and peace, the philosopher sits atop a hierarchy of less-valued behaviors and less well-behaved values stemming from the passions—this goal is potentially open democratically to all human beings.³²

Lenz's argument is vital for yet another reason. He demonstrates that we should not be limited to simply one course of knowledge. My own argument, I believe, is even more all-encompassing. Yes, knowledge, and ultimately wisdom, is supremely important, but whence come knowledge and wisdom? If we are to acknowledge knowledge, wisdom, and the exercise of those faculties comprising rationality, as we should acknowledge them, we must realize that we cannot conceptualize these supreme values simply as large, ivory towers full of books, professors, and students. I am, of course, only half right. We are all students, and we are all professors; we all have something to learn and something to teach. The classroom is the world of rational experience, and the books we take with us to the classroom, they are rational experiences. These experiences are rational, but we must continuously reinforce what we mean by this. These experiences can be highly emotional or, otherwise, deeply psychologically meaningful. Philosophers of language and analytic philosophers, more generally, often speak of literal meaning or emotive meaning, but there is scant mention in the literature of psychological meaning more generally, or even of the sort of meaning encompassed by the study of phenomenology. For instance, love is a rational experience in this way, although I believe strongly that it is not essentially an emotion; it is a rational, phenomenological concept.³³ Pleasure is also highly psychological, and more often than not, it is something like autonomously chosen prospective purposes which instill in us a sense of pleasure or happiness. In other words, it seems to be the case, and there is much psychological theory and research to reinforce this, that that person is happy who is free and autonomous and, conversely, that that person is not happy who is not free and autonomous; the greatest pleasure, it seems, cannot be derived otherwise. This, to be fair, was a point Mill drove home in *On Liberty*, although it was a fault of his utilitarian

³² Ibid., 96.

³³ Sophie Loidolt, "A Phenomenological Ethics of the Absolute Ought: Investigating Husserl's Unpublished Ethical Writings," *Ethics and Phenomenology*, eds. Mark Sanders and J. Jeremy Wisniewski. (Rowman and Littlefield, 2012): 1–37.

theory that he was unable to explicitly state the inherently greater value of freedom over pleasure.³⁴ It is, perhaps, for this reason that he made the fatal mistake of openly supporting British overseas imperialism. It is, perhaps, for the same reason that Bentham supported the brutality and inhumanity of the penitentiary system and its dreaded Panopticon that he, himself, designed; as it turns out, the question “Can they reason?” does, indeed, matter at least as much as “Can they suffer?”

6.2.10 Phronesis, Eudaimonia, and Nirvana

The Buddha supposedly stated that suffering is the result of desire. This, however, can and has been wildly misinterpreted. What is meant by this maxim is that all suffering is the result of unbridled, unselfconscious, and unrestrained desire, desire that is not modulated and regulated by the enlightened mind. The *telos* for Buddhism, of course, is *nirvana*, but it is not some state of blissful self-negation, as some have charged. It is, rather, a state of peacefulness with the world, a state of appreciation, a state of acknowledgment of one’s interconnectedness with the rest of existence, and a state of true self-knowledge and mastery over oneself. It is, in short, the participation in a particular, enlightened phenomenological paradigm. Suffering also connotes something very different in Buddhism than it does in Western cultures. It is not, Buddhist theorist Damien Keown has explained, a utilitarian notion of pain, a high amount of dolors, or a low amount of utils; it is, rather, a notion of dissatisfaction, of a lack of self-fulfillment.³⁵ As the Dalai Lama has stated concerning ethics, “mental peace” is the primary goal of Buddhism, and compassion is one prime key to such mental peace.³⁶ As he states:

Any action with certain purpose, carried in a realistic way, leads to a desired purpose fulfilled. Unrealistic approaches often fail to bring a result. Therefore, in order to carry a realistic approach, it is very important to know reality. An agitated mind cannot see the reality. When people have anger, then the object which they feel angry about, that object appears negative. . . . Ninety percent of that negativeness is mental projection. The negativeness is much exaggerated. Similarly, the positiveness of the object is exaggerated. So, with that, you cannot see the reality. So, with that, any sort of method, any approach becomes unrealistic. So there also—calm mind, balanced mind. Look at the object from various ways, from different angles. . . . Then you can see the reality, and then any action becomes realistic action. So, compassion brings calm mind. That helps to, firstly, open our heart to others. Then, a calm mind gives our brain the ability to judge reality. . . . From this, you can see the whole world as one entity.

³⁴ Mill, *On Liberty*.

³⁵ Damien Keown, *The Nature of Buddhist Ethics*, (New York: Palgrave. 2001).

³⁶ His Holiness the XIV Dalai Lama. *Ethics for our Time*. University of California Television. September 2009.

Constitutive of mental peace for Buddhism, therefore, is *phronesis*. The path to mental peace can be obstructed by many things, including desires derived from this exaggerated negative and positive “agitation.” The desire of which we are to rid ourselves is that desire connected to material pursuits, most notably money. We are to realize that, in this world, we need material things to live, but we must not get into fear or worry about circumstances beyond our control. Material things are all fleeting and impermanent; we will lose our health, we will lose our wealth, if we ever had any, but we can never lose mental peace if we pursue the tools (or virtues) necessary to achieve it. According to Buddhism, it is these virtues, and not material pursuits, that are truly valuable. Many people discover this only in old age, when they begin slowly to lose that material thing they are most attached to: their bodies. Others have mid-life crises. Yet others never come to realize this truth and, as a result, lead perpetually miserable and strife-ridden lives. Our state of mental peace is to be independent of material circumstances. We must realize that, yes, material pleasures and trinkets are often good in some way, but that those things will never fulfill our *teloi*; those things are nice and pleasant, but we should have the same courage, fortitude, calm, and balance after it is all taken away. Thus, one can see the Buddhist maxim “all suffering is caused by desire” more clearly: those who suffer when their material objects and pursuits have been taken away do so because all of their self-fulfillment has been made dependent on nearly worthless material things such as physical appearance, appetite for money, fine food, fine clothing, and other similar distractions. One cannot fulfill oneself through the material, only through the immaterial. Mental peace, in Buddhism, should then be completely independent of these states. When your house is foreclosed on, do you retain mental peace? When you lose someone you love, do you retain mental peace? When you are diagnosed with AIDS, cancer, or fibromyalgia, do you retain mental peace? Do you retain it even when you are being burned alive? This last example comes directly from the great mental peace many Buddhist monks in Vietnam during the brutal Ngo Dinh Diem’s reign exhibited when they protested his rule by setting themselves aflame—while meditating. They continued to meditate in total tranquility, on fire, until they died, and such self-immolations occur in Tibet to this day in protest of the Chinese government’s occupation. The point here is not asceticism. Indeed, when Siddhartha founded Buddhism, one of his prime motivations for doing so was to discover an *alternative* to asceticism. Enjoyment of material things is not bad; it simply can never be the source of self-fulfillment. As Epicurus and his followers warned in the stone inscriptions which they placed in the marketplace for all shoppers to see, material things can never bring happiness. In Gewirth’s own writings on self-fulfillment, he comes close, in some respects, to the Buddhist or, otherwise, virtue-theoretical view, but he is limited by his deontology. In rejecting the Aristotelian notion of natural ends, he is misled, I believe, back down the path of desire (and perhaps even internalism), in the form of what he calls “aspiration fulfillment,” as a legitimately guiding force for one’s fulfillment.³⁷

³⁷ Gewirth, *Self-Fulfillment*.

These desiderative conceptions of self-fulfillment, we must realize, are ultimately vapid, for they do not invoke the concept of *telos*. Indeed, the concept of *telos* is conspicuously absent even from many accounts of virtue ethics, including the Dalai Lama's conception of Buddhism. Mental peace, for instance, should not simply be an end in itself. Practicing and attempting to attain mental peace by itself is equally vapid. Mental peace, as the Dalai Lama himself aptly points out, is key, as he puts it, to "seeing the reality." But what is "seeing the reality" for? Why *should* I wish to see the reality? The answer, we must see, is that I must work toward this state of mental peace and clarity, this state of accurate and realistic perception and self-perception, precisely *for the purpose* of striving toward my end. It is precisely the end, the *telos*, which gives one's life purpose and meaning, and as such, the focus on *striving* should be our principal aim. This concept of striving is important, for when I speak of striving, I imply a state of being in which one's consistent and perennial conscious focus is upon one's *telos*. Even in the lowest of material or psychological states, our sights should be insistently and unceasingly set upon the kind of person we must be, even if the attainment of that ideal state is improbable. Some may think of this as unrealistic, but it is indeed the most realistic approach conceivable. Whatever situation I am in, if I am not to sink even lower and lower into horrifying depression or material squalor, I must make a commitment to myself to survive. What I mean by this is not only literal survival, but also survival in the sense that I must keep moving forward despite everything and achieve this by keeping my sights set upon my *telos* or purpose. As it is *a priori*, all purpose is perpetual and indelible; it can never be destroyed, only forgotten. Once I have made this commitment to myself, I must then *act*, keep on acting, and never look back. So long as I continue to strive, and so long as I continuously focus upon striving, I will know that I am a greater and more excellent person than I was yesterday; I will know that, even if I die tomorrow, I became the manifestation of my meaning, and I have fulfilled more meaning than perhaps the most materially wealthy, most psychologically sound, and most superficially happy individual.

Yes, desires are significant, but, as in Buddhist, Epicurean, and Aristotelian ethics, we must ensure that our desires are good in that they fulfill our *telos*, whether it be called *nirvana*, *ataraxia*, or *eudaimonia*. In order to discover which things are our individual *teloi*, we must not, of course, do evil to others or, otherwise, neglect others. Doing good for others is, in fact, a vital part of our collective *telos*, and if it is a part of our collective *telos* as agents, then this function of doing good for others is also part of our individual *teloi*; indeed, it is an essential part of my, your, and her individual identity. We have discovered, then, that, whatever a person's individual *telos*, that is whatever specific function or activity it is that constitutes his specific rational end, included as part of that *telos* must be to do good for others. Doing good for others, in other words, generally coincides with doing good for oneself and vice versa. This is because, in striving to reach out to others and give to them in some way, we aid them in fulfilling their own *telos*, which makes them a better, more excellent, and more enlightened person; this, we must realize, is consistent with working toward the fulfillment of rationality's *telos*, since it is an action which furthers an individual's movement toward the ultimate goal of the best exercise of rationality. As a result of such an action, we instill virtues in ourselves in the process.

6.2.11 *The Question of Individual Telos*

We have not, however, yet answered the question of what different people's *teloi* are; we have not determined what every single person should do, whether it is being an astrophysicist, a mechanic, a street performer, a librarian, or a concert pianist. The simple answer, if we remember that rational experience, and not simply one specific setting, is our classroom, is that it is up to all of these individuals, themselves, to ascertain their *telos*. We can all help others in discovering their *telos*, and indeed we should, but, in the end, it is that person alone who can ascertain her *telos*. Why is this? First, different people have different strengths and weaknesses and different modes of *techne* to which they find themselves particularly suited. In other words, everyone has different ways in which they are physically, mentally, or otherwise capable of exercising their rationality well. If I can make beautiful grandfather clocks and I am very good at making grandfather clocks, which themselves necessitate a great exercise of rationality, yet I cannot play piano well, or at least I cannot play piano well enough to become a concert pianist, it is better for me to pursue, as one of my life goals, the creation of new, excellent, and artistic grandfather clocks. This, of course, does not disbar this person from playing piano if she or he wants to; it simply disbars him from ever doing it truly excellently. She or he may be the kind of person, of course, who can do both, in which case she or he will discover, on her or his own, which purpose it is best for her or him to pursue. This, however, it bears repeating, is something that an individual must ascertain, although it is important for others to help this person discover her or his *telos*. This process is a shared experience. As Macintyre has stated, "Genuine and extensive self-knowledge becomes possible only in consequence of those social relationships which on occasion provide badly needed correction for our own judgments. When adequate self-knowledge is achieved, it is always a shared achievement."³⁸ This also does not mean that this person should not seek knowledge and wisdom through all possible areas of academic study, including biology, chemistry, physics, poetry, literature, drama, music, mathematics, anthropology, sociology, history, philosophy, and others; it simply means that it is not the case that we should expect every single person to become an expert in or become excellent at every single possible field of study. Thus, we see that everyone has a rational *telos*, but that every individual has at least a slightly different, and sometimes a wildly different, *telos*. We see, also, that there are certain virtues derived from the *summum bonum* which are so derived because of their necessity or tendential necessity for the *summum bonum*. These virtues are those which, whatever *telos* we discover that we have and whatever we do in our lives, we must cultivate and exercise. Taken from our previous analyses of the faculties of best rationality, and alluded to somewhat in those analyses, we can now see that, in order to be an excellent or enlightened person, we must cultivate at least the following virtues: self-awareness, self-control, contemplation, *phronesis*, creativity and

³⁸ Macintyre, *Dependent Rational Animals*, 95.

imagination, compassion, and justice. The experiential conception of rationality, comprised of conceptual synthesis, conceptual abstraction, and freedom, elucidates the importance and salience of these virtues, and this is one of the reasons why I have highlighted these features of rationality in addition to the conception of logical rationality that is typically emphasized in deontological and even many virtue theories. Experiential rationality, along with logical rationality of course, is what makes such *ergoi* as art and aesthetic experience, in general, the goods that they are. Experiential rationality is also responsible for the goodness of such experiences as smelling incense or essential oils, seeing a landscape during sunset, hiking up a forested mountain, meditating, or hearing a Mozart symphony or soul music. Granted, it is not the case that all people will experience these same experiences in the same or even in relevantly similar ways, and some people will dislike the smell of certain incenses, but all of these other activities, whether or not they invoke in us a sense of pleasure at first, appeal to and make more excellent our experiential rationality. They achieve this, partially, by drawing upon concepts from the experiences in our memory and thereby constituting for us a part of the narrative that constitutes our lives and our individual essences. St. Augustine was, and still is, one of the very few philosophers who cited the ethical significance of memory,³⁹ and, as one of the major sources of our ability to abstract and synthesize concepts, its role and significance in individual *telos* should be more comprehensively explored. Parfit's claims about personal identity and psychological continuity, if true, would not undermine the essence-accident distinction with regards to memory, personal identity, and individual *telos*, for if his "relation R" is what matters, and if relation R is something intelligible and extant, then it must have an essence, no matter how complex, as well as a corresponding *telos*. This, however, is a far more complex debate and discussion than can occur substantively here in this book.

6.2.12 *Some Implications for Aesthetics*

The importance of the phenomenon of memory, as a fundamental component of personal identity, is perhaps why the appreciation of music and other arts should start at a young age. I, as the son of a vocal pedagogue, have been fortunate enough to have grown up around music and to have been exposed to music for much of my waking life. The appreciation of great music, therefore, has been a part of my own identity for as long as I can remember, and, thus, I observe how difficult it sometimes is for others to hear the nuances, subtle beauty, and complex underlying patterns of such music. Thus is it that people who have never listened to such music or seen great visual art are rarely ready, from the start, to appreciate it; they often have a difficult time hearing these nuances or feeling the message that the artist is communicating. Why is this? This is because the best art requires certain virtues to begin with, namely

³⁹ Augustine, *Confessions*, (Oxford: Oxford University Press. 1998).

patience, concentration, and engagement. While that art which is bad is so marked by its tendency to leave the person a passive spectator (see Adorno and Horkheimer's critique of mass culture), it is good art which demands a response, a response characterized by the subject's active role in the art. In order to be good art, in other words, the art must appeal to these rational activities by demanding that the subject become almost meditative, concentrating deeply on the nuances and interweaving complexities, and often the simplicities, of the art. Listen to Bach's "Harpsichord Concerto No. 1 in D minor (BWV 1052)," Strauss's "Zueignung," or the final scene of Puccini's "Suor Angelica," (particularly Diana Soviero's rendition) but concentrate on it; focus all of your attention on it; meditate on it. At first, listen to the voice part. What has the composer done with the voice, and why has he made the vocalist use certain phrasing, changes in tempi, dynamics, staccati, or a certain melody or recitative? Then, listen once more. Now, listen for the blending of the instruments. Why has the composer done it this way, and what effect does it have? Do not listen to it again for a few hours or even a few days. Once you listen again, listen to something else, perhaps the chord progression itself or simply one particular instrument. Perhaps, if the main part is the vocal line, and it is an aria from an opera, you might try to empathize with the character or even "become" the character by attempting to adopt his or her phenomenology. One can use other conceptual frameworks from such fields as film theory, for instance, to analyze film in much the same way: what are the mise-en-scene elements, what if any are the signature elements of the auteur, what are the motifs, and how is all of this expressed? This is why such art is so rich and so good; you can keep going back each time and learn or discover something new. This is the only way art can be fully appreciated: by reflection via conceptual abstraction and analysis. Good art, I must also stress, is not limited to a particular genre or set of genres.

What, then, does all of this reveal about the value of art and aesthetic experience in general? I will not, for the sake of space and the reader's sanity, undertake a comprehensive analysis of the implications of my theory for aesthetics. Instead, I will merely state a few observations. First, we must see that aesthetic experience, generally, must possess high value, for it seems highly implausible that a person will be able to achieve her or his respective *telos* without an exposure to various kinds of art and aesthetic expression. Art, then, seems to achieve its value solely from its instrumentality for the fulfillment of the individual's *telos*. I believe the application is broader, however, than the mere experience of art in a formal setting; I believe, rather, that the theory demonstrates that potentially every experience is an aesthetic experience and that, as such, the individual should thus attempt to draw out the aesthetic experience inherent in everyday life. Thus, the value of art truly finds its value in the overall value of phenomenology, a topic we shall also discuss.

6.2.13 The Question of Individual Telos Revisited: The Relevance of Phenomenology

I understand I have gone off on the subject of art and aesthetics, but I believe it is important, for one of the main points of this book has been to demonstrate that

analyses and derivations of ethics and value from the phenomenon of logic or logical rationality alone are futile. Yes, they can yield much in the way of firm theoretical foundations, but because they largely ignore the significance of experiential rationality, they miss half of the picture and have certain gaps as a result. One might even accuse me of thus far leaving out the richness that is experiential rationality or of undertaking an analysis that is still too dry and “logical” to capture the idea that is experiential rationality. This is, in some ways, true. I have thus far focused merely on the empirical aspect of experiential rationality, but I have not touched at all on the phenomenological aspects. Experiential rationality, while indeed empirical, is also overwhelmingly phenomenological. When we observe and understand, we first do so through the internalization of meaning. When we see a desk, a bed, a chair, or a tree, these objects are not merely objects of *Verstand*, that is, they are not objects that can be described solely in terms of the material of which they are made, the configurations which comprise them, or the physical interactions they have with other objects. Each of these objects, by contrast, are, from a phenomenological perspective, also “made of meaning” and, just as we might break down objects and reduce them to their constituent physical parts and characteristics, phenomenology aids us, mostly by way of hermeneutics, to break down these objects into their constituent meanings. When this process takes place in writing or speech, we call it narrative. Even narrative, furthermore, has an underlying logic. I have emphasized thus far the importance of analyzing and discovering one’s personal identity. A study of empirical reality, while helpful, will not however do justice to such analysis and will leave much neglected from such discovery. One’s personal identity just is one’s essence, and from this essence, one can derive one’s *telos*. Furthermore, without a personal essence and *telos*, the faculties of rationality have little determinate content. The faculties, rather, merely constitute the formal framework of excellence. It is personal essence which gives these faculties determinate content. Thus, this is why a vital part of the *summum bonum* consists in self-discovery, self-analysis, and self-knowledge. If identity is largely, if not fully, phenomenological, then it follows that phenomenology plays a large role, beyond that of empirical psychology, in the constitution of one’s essence and subsequently derived *telos*. This, however, raises another important question: “How, given the phenomenological framework within which I exist, do I determine my essence and then use the faculties of best rationality in order to determine my *telos*?” First, as I have stated before, the process of self-discovery is an individual process, not because I want it to be that way, but because it necessarily just is that way. I cannot peer into your mind and experience the phenomenological paradigm that you do. As such, I can only aid you in discovering yourself by being your friend, by interacting with you, and by engaging in discourse and other communicative action with you. In the end, however, it is only you who can discover your essence and, from this, proceed to discover your *telos*. “From the discovery of my essence,” one might ask, “how do I then proceed to discover my *telos*?” One’s essence contains the answer. Let us notice first, however, something interesting about the faculties of best rationality. In their generality, the faculties of best rationality do not have much content. That is because this formalistic context necessitates an essence. Once such an essence is given, the *telos* can be found.

Think, for instance, of all of the phenomenological and empirical components that comprise hypothetical person A. Person A has a particular history, a particular memory, and from this, a particular fundamental phenomenological paradigm which at least partially constitutes her essence. If applied in a phenomenological manner, instead of in a merely psychological manner, we might be able to call this essence by Parfit's "relation R." This relation R constitutes the underlying logic which comprises the relations between experiential objects x , y , and z , as well as between experiential events x' , y' , and z' . This underlying logic then comprises one's identity. Of course, the list of experiential objects and events would be extremely far more comprehensive than this, but this example is still illustrative of one of my central points, namely that one's identity cannot be fully captured through the empirical science of psychology, for who one is not fully a scientific question. It is, rather, a question of meaning or meaning-relation, and where in the *empirical* world do we ever *observe* meaning? This leads to my next set of claims. In order to derive a *telos* from this information, person A needs to engage in self-reflection. In other words, in order even to begin discovering how freedom, conceptual synthesis, conceptual abstraction, and logical judgment best manifests in herself, she must engage in self-reflection, self-analysis, and, as Habermas importantly emphasizes, interaction and communicative action. As McCarthy points out with regards to Habermas's treatment of the subject: "Knowing himself to be involved in his development, to be a result of the 'history of consciousness in its manifestations' on which he reflects, he must direct the critique of ideology at himself. In this way critical theory pursues self-reflection out of an interest in self-emancipation."⁴⁰ On my interpretation of this, it is such self-critique which allows person A to eschew false judgments, not simply about ethics, but also about what the world is like. Critique, including self-critique, can liberate person A by leading her to question her own assumptions about what the world is really like. From this move, she advances closer to her *telos*, for she moves closer to practicing best logical judgment which, I have shown beforehand in various proofs, is equivalent to correct and accurate logical judgment. This move thus changes x , y , z , and x' , y' , and z' and begins to shift them toward a different phenomenological paradigm. Through critique, these phenomenological objects and events are made better. As McCarthy continues:

It is this connection between critical reflection and liberation from 'dependence on hypostasized powers,' from 'seemingly natural constraints,' that underlay Habermas's 'fourth thesis' in his inaugural lecture at Frankfurt University: 'In self-reflection knowledge for the sake of knowledge attains congruence with the interest in autonomy and responsibility.... In the power of self-reflection knowledge and interest are one.' Since critical reflection undermines the dogmatic character of both a world view and a form of life, the cognitive process coincides with a self-formative process: knowing and acting are fused in a single act.⁴¹

⁴⁰ Thomas McCarthy, *The Critical Theory of Jürgen Habermas*. (Cambridge: The MIT Press, 1978): 88.

⁴¹ *Ibid.*, 88.

Importantly, interaction is integral to this process as well. It is extremely difficult, if not impossible, to gain self-knowledge concerning who and what we are without engaging in communicative action with others who can help us. While I gave examples of *teloi* earlier that were what Habermas would call work-oriented, interaction, as a fundamental part of self-discovery and self-knowledge, is also a vital part of person A's *telos*. Such phenomenologically important notions as friendship and love are thus a part of person A's *telos*; she indeed cannot achieve excellence without such friendship and love and the meaning-relations which accompany such phenomenological paradigms. Thus, person A's *telos* is not, in common parlance, only about "what one does;" it is also, for lack of a better phrase, about *how one is*, or the state of one's interaction-oriented being. At this point, person A must take a good look at the changed and changing underlying logic between x , y , z and x' , y' , and z' and determine, from these phenomenological relationships, what manifestation of freedom is possible for herself and which, furthermore, is unique to herself. "What kind of possible meaning," she must ask herself, "is entailed by the underlying logic of the meaning-relations which comprises my phenomenological essence?" Once again, this is an extremely difficult question to answer, and as such is a highly personal question. Person A, we must see, is ultimately the best judge of this. Once again, however, we see that logical judgment is important here. If to judge correctly and accurately is *phronesis* or best logical judgment, then *phronesis*, when it is applied to the analysis of the self's constitution, reveals the underlying logic of the meaning relations which comprise one's phenomenology and thus is necessary for the discovery of one's essence and, thus, also one's *telos*. We see this integral nature of logical judgment as applied to phenomenology when we consider that such a statement as "If best logical judgment is best for the discovery of one's essence, and if the discovery of one's essence is necessary for the discovery of one's *telos* and the manifestation of best freedom, best conceptual synthesis, and best conceptual abstraction in accordance with that *telos*, then best logical judgment is best for the discovery of one's *telos* and the manifestation of best freedom, best conceptual synthesis, and best conceptual abstraction in accordance with that *telos*" is *a priori* true. It follows from this that best logical judgment or, as I have analyzed it even outside the context of these other faculties, is both part of and absolutely instrumental to the discovery of one's *telos*. From this move on, the determination of person A's *telos* by me is probably impossible, since the number of possible manifestations of freedom, and thus of liberation, is probably impossible to count and delineate with any degree of accuracy. Person A is ultimately the only one who can discover the shape, nature, and ultimate meaning of her *telos*. What we do know, however, is *how* her *telos* will be delineated. We do know, because of the logical relationships we have previously traced between logical judgment and the other faculties, that such a *telos* will necessitate the virtues of self-critique, self-analysis, self-awareness, self-control, and self-re-conceptualization. These virtues are most closely associated with best freedom. The virtues of contemplation, *phronesis*, and learning might also be added here, although they are clearly also very closely associated with both best conceptual abstraction and best conceptual synthesis. The virtues of creativity and

imagination, of course, are also integral to the excellence of these faculties. The virtue of justice is also encompassed by *phronesis* or best logical judgment, as justice is the ability to make sound judgments concerning what is right and wrong with regards to the treatment of others and oneself. Indeed, as a basic truth of this virtue-theoretical system, it can be seen that it is a manifest contravention against sound judgment to act in ways that degrade, devalue, or otherwise hinder the excellence of other rational beings. This can be seen in the simple statement: “If best rationality is the *summum bonum*, then it is evil to hinder the development of best rationality.” The theories of Habermas and Gewirth reaffirm this truth.

6.2.14 *Gewirth and Habermas Revisited*

At this point, we must admit that Gewirth and Habermas’s respective theories have yielded much in the way of firm theoretical foundations, specifically in the area of rights, duties, and justice. Their theories, we must see, however, are ultimately dependent on virtue theory both for their optimal applicability and for the intelligibility of the axiological terminology employed therein. This, once again, is because their theories demonstrate to us that we must act in certain determinate ways, and that it is unjustifiable to act in certain determinate ways, but all of this is dependent on the fact that we, as agents, must make certain claims to freedom and well-being. We must make these claims, but, indeed, our making these claims does not make them true. A dedicated Gewirthian would tell us that this is not a problem and, to a certain extent, it is not; strictly speaking, it is not necessary to demonstrate that “X is good” or “X is better than Y” are true in order to prove that all agents have equal rights to freedom and well-being. This is not the issue. The issue is that there seems to be something very wrong with a theory which shows us what we must all claim and, thus, what we must accept in our interactions with others, but which does not demonstrate to us what these claims *mean*. As a result of dialectical deontology’s dependence on the existence of other people for its legitimacy, it suffers from what I have previously called the Single-Person Problem. By contrast, my theory of the good, and my subsequent interpretation of virtue theory, suffers from no such self-other asymmetry problem. On my account, the existence and intrinsicality of goodness is not dependent on the instantiation of a relationship, logical or otherwise, between two or more people; neither is what I have called betterness or bestness. My theory of the good is an autonomous, deontology-independent one. There are certain things, we have shown, that are better for people to do and to refrain from doing. To aid someone in distress or poverty, *ceteris paribus*, is to be done because it is virtuous; it has a tendency to lead to the cultivation of that person’s virtues, and, subsequently, has a tendency to lead to the cultivation of our own. This, it can be shown, is better to do, *ceteris paribus*, because it is tendentially necessary or good for the excellence of rationality. Since we have shown that such an act, among myriad others, is, *ceteris paribus*, better than not doing the act, then, because of what we have discovered concerning the

structure of value judgments and judgments of goodness, we can conclude that we ought to do it or, if both possibility and permissibility are met, that we must do it.

The reader should note this as significant. Why? We have just derived the deontic terms “must” and “ought,” as well, implicitly, as “should” from a judgment of goodness and of what is virtuous. Such deontic terms, we should note, cannot derive judgments of goodness. What I mean by this is not that these types of judgments do not have a bi-conditional relationship to each other, which they do. Instead, what I mean to say is that we cannot, if presented solely with a deontic judgment, make any judgments whatsoever concerning what is good, since, as we have shown, the meaning of such deontic terms is dependent upon the structure of value judgments and judgments of goodness, and the structure of such judgments is, in turn, dependent on the meaning of word “good;” the phenomenon of choice, itself, is dependent, first and foremost, on the presence of judgments concerning what is better or best and not, strictly speaking, what is categorically obligatory. Categorical obligation is, rather, derived from an initial recognition of what is good, better, or best. Through our definition of goodness, surely, we have seen that deontic concepts, especially that of necessity, are essential to goodness, but this is somewhat beside the point. The fact is that it is impossible to know what one should do or must do without, first, discovering what it is better, best, and most virtuous to do; this is simply the pattern of the practical, although not the theoretical, “must”’s logical derivation.

Thus, in addition to Gewirth and Habermas’s respective dialectical methods, we have established an assertoric method which is consistent with both the Principle of Generic Consistency and with the conditions of the Ideal Speech Situation. How do we know that my assertoric theory of the good is consistent with Gewirth’s PGC and Habermas’s principles (D) and (U)? This is simple. We know, for instance, that Gewirth’s entire theory rests on the first dialectical premise “I do X for purpose E.” As we have seen, and as Gewirth’s theory corroborates, if I make this assertion, then I must also make the assertion “X is better than Y,” although Gewirth makes this same point by offering us the assertion “X is good.” We know, from here, how this works. Eventually, this results in the statement “All agents make right-claims to freedom and well-being.” Thus, from the perspective of all agents, they all have certain determinate positive and negative rights against and duties to each other. My theory of the good, however, now gives these premises *assertoric* content. Instead of stating, dialectically, that all agents assert “X is better than Y,” we can actually make a true or false statement concerning X’s being better or worse than Y, and we can prove it to be true or false. From here, we can demonstrate, assertorically, that one ought to or must do X rather than Y. Also assertorically, we can assert, “Agents must be free to do X rather than Y” and “Agents must have the freedom to do X rather than Y.” Thus, we can assert, also assertorically, “Agents must have the freedom to pursue some good X.” At this point, one might object that, if one ought to have the freedom to do only ethically justifiable or good things, then one might as well set up a Stalinist regime and force people to do only good, virtuous, ethically justifiable things all the time. Let us be very careful, however. By asserting “Agents must have the freedom to pursue some good X,” let us be wary of what we mean. By no means does any such regime or order follow from my theory of the good. Why? According to my account,

as is the case with Gewirth's account, the use of any more force than is necessary in order to stop some bad action is bad and, thus, vicious. Furthermore, and more importantly, it is bad, and therefore vicious, to halt or prevent some evil using some greater evil. To outlaw smoking, for instance, is vicious because it denies people the freedom and autonomy they have to make their own choices, and such a forcible limitation and suppression of choice, we have seen, is vicious. Yes, smoking is a vice since it has the high tendency to make one's life less excellent, but the forcible suppression of smoking by a centralized power, and even by de-centralized powers, is, we must recognize, a suppression of choice and autonomy. Other steps that could be taken include policies to ameliorate the suffering of the poor and thus perhaps lessen the tendency to turn to smoking, to contribute to consumer education, and to provide better access to health care, health information, and psychological services, the latter which is especially salient for addressing why many people turn to smoking in the first place. A forcible treatment of agency in order to solve this problem, we must see, is vicious. Such forcible treatment of agency, indeed, undermines the very principles we are trying to affirm and uphold.

Let us return to Gewirth's proof. "Agents must have the freedom to pursue some good X," we said, can now be established assertorically. Since it can be so established, as we have seen from our analysis of "must" and its logical relationship to rights and duties in Chap. 4, we can also establish, assertorically, that agents have the right to pursue some good X. As we can see from Gewirth's theory, this is also the case. Agents have the right to pursue goods, but *others do not have the right to force or coerce others to pursue such goods if a greater evil, namely a hindrance of freedom, will result from the use of such force or coercion*. In other words, this has the same practical consequences as the Principle of Generic Consistency, which, it helps to repeat, states: *Act in accord with the generic rights of your recipients as well as of yourself*. In other words, contrary to Gewirth's assertion that rights are necessary for the establishment of virtues, it seems that, assertorically, virtues are necessary for the establishment of rights. As for Habermas's theory, it suffices to say that, since we have previously established that Habermas's theory is consistent with Gewirth's theory ($H \leftrightarrow G$), and since we have also shown the aretaic value theory I have offered to be consistent with Gewirth's theory ($A \rightarrow G$), then it follows that the aretaic value theory I have offered is also consistent with Habermas's theory ($A \rightarrow H$).

6.2.15 *The Single-Person Problem Resolved*

Virtues, however, are necessary for more than an assertoric establishment of deontological theory; they are, due to the Single-Person Problem, necessary for the establishment of any comprehensive deontology, including Gewirth's theory. In the single-person world, we might recall, no deontological, consequentialist, or other rule-based theories have any application. This, again, is because all of these kinds of theories are predicated upon certain logical or other relationships

instantiated by the existence of other people or, at least, other sentient beings. Thus, even Gewirth and Habermas's theories fail to solve the problem because their basic structure is also predicated on the logical structure of justifiable and unjustifiable interpersonal relations. It says nothing, for instance, about what I must do or what it is better that I do; instead, it only tells us what I must consider good and what kinds of "must"-judgments I must make. Thus, we see the problematic nature of the dialectical method in the Single-Person Problem, namely that, in the absence of other agents or even other sentient beings, my necessary "must"-judgments and judgments of goodness provide me with no behavioral directives; they only provide me with the information that there are certain things that I must consider good. My theory of the good, and virtue theory in general, holds the solution. We have already established that best rationality is the *summum bonum*. Since life is tendentially necessary for rationality, I can deduce that it is virtuous to uphold my life. This is, then, my first priority. This means I need food, water, and shelter. Let us assume, for the sake of argument, that it is easy for me to obtain these things in all the abundance I need. Then what? As we previously established, since I am at an adequate enough level of Maslovian need-fulfillment, I must go on a search, so to speak, for my *telos*. Let us be reminded that a single-person world is a truly tragic world, if only for the reason that no discourse between people can possibly be had to aid this person in his or her search for his or her *telos*. We do not have to go too far with this. The point is simply that, even in a world where there is but one person living, additionally without any sentient beings, my theory of the good and virtue theory, in general, is capable of far-reaching applicability.

6.2.16 *The Final Assessment: Choosing a Model*

Thus, we have shown, virtue theory seems to be the most explanatorily powerful account of value theory. Virtue theory answers a number of questions that deontological and other rule-based theories fail to, including but not limited to: "What does 'good' mean, and what is goodness?," "What are my meaning and purpose in life?," and all of the questions mentioned at the beginning of this book which I believe derive from these two fundamental questions. This is great for virtue theory, but something now is staring us in the face: we have three very viable theories. Which one, therefore, should we choose? Should we choose Gewirth's account, Habermas's account, or mine? This is a trick question, for the answer is all three. All three, we must see, have the same practical result in the context of society or alterity generally. As I have argued previously, Gewirth and Habermas's theories are essentially mirror images of each other; the syllogistic structures of their respective theories commence on different premises that each theory nonetheless shares at some step in the syllogism, yet each theory ends up with essentially the same conclusion. They have, therefore, a sort of functional fit with my own theory in that they reinforce the faculty of logical judgment and thus serve as sorts of

modules or tools for this faculty to employ in its judgments and thus in the pursuit of its own excellence.

We must recognize, however, that there are some seriously substantial differences between the three theories. Gewirth's, for instance, despite his long treatise on self-fulfillment, still partially bases its account of self-fulfillment on desire, a potential mistake which may teeter his theory of self-fulfillment into internalism. This predication of self-fulfillment, even partially, on non-rational or first-order or even many second-order desires, however, is a feature of many rule-based theories which such theorists as Foot, Geach, Anscombe, and Macintyre have notably critiqued. Habermas, too, seems to devalue value judgments themselves, for much of his analysis tends to depict value judgment as a largely emotional phenomenon. My theory of the good and account of virtue theory, however, does not fall into these kinds of traps. Not only have I provided a definition of goodness and an account of intrinsic goodness which fits Magnell's rigorous criteria, but I have also provided an account of what I call betterness and bestness, and I have shown that these variants of goodness are what I have called *reflexively intrinsic*. Additionally, I have established that the *summum bonum* is best rationality or wisdom, which I have defined, and I have demonstrated that, from this discovery, we can form directives concerning what is virtuous and what is vicious. Furthermore, I have done all of this without predicating any of my premises on agent desires. This is significant, for it demonstrates that any virtue-based theory of self-fulfillment will be, indeed, far more profound than that which any rule-based theory can give us. Virtue theory, for instance, offers us insights concerning how it is better to treat ourselves, how we are to conceive of and come to terms with our personal identity, and, ultimately, how we are to conceive of our place in the empirical and phenomenological universe. This, I believe, is momentous.

6.3 Conclusion

6.3.1 *The Eternal Nature of the Summum Bonum*

There are theorists who have reasoned that, if there are laws of physics, laws of mathematics, and laws of logic, it is not out of the question to posit laws of morality or ethics. Indeed, natural-law theorists took and still take this position with regards to ethics. The positive-law tradition, of course, overshadowed the natural-law tradition in Europe in the nineteenth century, with its supposed pragmatism and practicality. This pragmatism and practicality, of course, along with other so-called pragmatisms, such as Bismarckian *realpolitik*, exposed themselves throughout the century and beyond as mere prejudices conducive only to horror and destruction. This is because legal positivism is actually a euphemism for relativism or subjectivism, both of which, when injected into a culture, must inevitably end in such horror and destruction. Positivism and relativism, we must finally admit, are neither pragmatic nor practical,

and, although theorizing has been criticized for its impracticality, we must see that it is actually far more practical than any legal positivism or ethical relativism. Maybe, however, the positivists were right about one thing: theorizing, alone, does not achieve results. Let us not, then, drop theorizing altogether, which was the terrible mistake of positivism; let us, rather, combine theory with action. “Why?,” we should ask, “are relativism and subjectivism so dangerous?” The answer is simple: their number-one danger is in their promotion of complacency. If nobody’s ethical claims are right or wrong, then there can be no legitimate criticism of any particular act. If Stalin thinks it is okay to imprison, torture, and execute supposed Trotskyites, or if someone of a particular group wants to subject someone else to genital or other bodily mutilation against her will, or if a government wants to commit car bombings and other terrorist attacks against civilians, or if Lot wants to submit his daughters to a mob to be raped, or if someone wants to bind a woman’s feet, or if someone wants to enslave someone else because she or he is black, or if someone’s religious beliefs impel her to despise all LGBTQ people, or if someone wants simply to tell a series of lies to cheat someone else out of the money he or she needs to live, then there is no legitimate criticism we can make of any of these beliefs or actions. We get stuck in the social rut of making excuses for the evils in people’s cultural practices, religions, or overall views, which include, but are not limited to, such excuses as “But it is their culture,” “It is their religion,” or “It is the way he wants to live his life.” This, of course, does not give us an excuse, for instance, to use force or violence in order to curtail such evil. In previous chapters, the reader has gotten a glimpse at those rare times when such force is justifiable and even obligatory. In most cases, however, it is discourse or mediation—not force—which is necessary. Especially as Habermas’s theory has demonstrated, it would be wrong to engage in violence or coercion in order to enforce our opinions upon others. In opposition to the relativist, we have, indeed, seen that there is objective goodness and even objective bestness and, furthermore, that the *summum bonum* is eternal. Philosophers often implicitly deride such a view. Once again, as Railton states: “The notions of good or bad have a place in the scheme of things only in virtue of facts about what matters or could matter to beings for whom it is possible that something matter. A being for whom something can matter is a being with a point of view, a subjectivity. In a universe without subjectivity, there is no value either.”⁴² This view, however, oversimplifies the issue. What do I mean when I say that the *summum bonum* is eternal? As mathematical laws and logical laws hold eternally, whether there are agents in the universe or not, the laws of morality or ethics hold eternally. Just as the relationship $((P \rightarrow Q) \rightarrow (\neg P \vee Q))$ holds whether or not there are agents, so does the relationship “Rationality is the most tendentially necessary for the teleologically comparative tendential necessity of functions” or “Rationality is the best for betterness” hold, whether or not there are agents. Few would make the bold claim, for instance, that, if all sapient beings in the universe went extinct, the earth would suddenly lose its circumference or that the trigonometric properties of radio or infrared waves would disappear. We do not, then, invent this

⁴² Railton, “Facts and Values,” 56.

value relationship; instead, we come to realize it. To compare something as better or worse than something else is to make a claim, not concerning some arbitrary *telos*, but concerning the essence and *telos* of betterness. Thus, if such relationships as $((P \rightarrow Q) \rightarrow (\neg P \vee Q))$ are the case, have always been the case, and always will be the case, then not only is rationality the *summum bonum*, but it always has been the *summum bonum*, and it always will be. Without this truth, it would be incoherent to state that, if there were no rational beings in the universe, the universe would be an abysmal and tragic place. This, indeed, is true.

6.3.2 *Enlightenment and Liberation*

The claim will inevitably arise, once again, that rationality's status as the *summum bonum* is elitist, and I believe it is, once more, important to tackle this concern. It is not the case, for instance, that everyone can achieve all of the goods necessary for best rationality. This, we must admit, is true. There are, and have been, oppressed people throughout the world who are in their condition because of discrimination, ignorance on the part of others, some form of elitism, or all three. This, however, does not change the fact that rationality is the *summum bonum*, and it certainly does not mark oppressed people as inferior. On the contrary, rationality's status as the *summum bonum* becomes, for us, a directive—a directive to contribute to the excellence of oppressed people by aiding them. Let us be emphatically clear, as well, that this is not some paternalistic directive, a directive which has, throughout history, more often resulted in a greater lack of freedom for its patients. Unlike with the imperialistic and paternalistic approaches of the past, I have pointed out that it is not simply better logical rationality that makes someone more excellent; it is, rather, logical rationality *and* experiential rationality, and only oppressed people, themselves, and all individual people in general, know their own experiences. We cannot tell people what they experience or how they experience. We must, rather, engage in discourse with oppressed people and try to understand their lives. Only if we listen can we understand, and only the individual herself or himself can find his or her rational *telos*. We can aid others in finding their rational *telos*, but, in the end, it is only they who can enlighten themselves.

We must not neglect, of course, the oppression of non-human animals and even the destruction and abuse of other forms of life. In the film “La Strada,” by Federico Fellini, the character, Il Matto, tells Gelsomina that everything has a purpose, “even this pebble. . . even the stars.” He was right, and our previous analyses have shown this to be the logical conclusion of a consideration of formal and final causality. We can help all living things achieve their *teloi* by living virtuously ourselves.

Rational beings, however, need more than survival for the fulfillment of their *teloi*.⁴³ They need certain virtues that are specifically rational ones: self-awareness,

⁴³ I do not here mean to imply that all non-rational living beings have only survival as their *telos*. In order to determine their respective *teloi*, of course, we would have to engage in serious inquiry into their respective essences.

self-knowledge, self-discipline, resiliency, creativity, compassion, and the ability to engage in discourse (broadly construed), among others. These virtues, we see, aid in yielding better individual autonomy and freedom, better conceptual synthesis, better conceptual abstraction, better judgment, and a better life overall. If these faculties of rationality are fulfilled in an individual to that individual's greatest capacity, then that individual is living the best life that she or he can possibly live. The excellence of rationality in an individual life, therefore, is both necessary and sufficient for what has traditionally been called "the good life."

Such philosophers as MacIntyre, Habermas, Apel, and Sen have stressed the importance of discourse in ethics, and they are right to do so. The importance of free and open discourse cannot be stressed enough. Indeed, it is due in great part to the lack of such discourse in a society that prejudices of various kinds exist. We live not in an open society but in an isolated society, enforced and perpetuated by what Habermas, Foucault, and others have exposed as institutionalized structures of power and domination. Because of such hypostasized powers and power structures as exist in the political system, the legal system, the financial system, the labor system, and in the professions, discursive and cooperative incentives have become marginalized. It is perhaps because of such power structures that, upon a superficial analysis of society, our interests seem so disparate. But, as we have seen, my interests cannot be fulfilled without yours also being fulfilled; it is fallacious for me to claim that I have become excellent if, in every step I take, I have not advanced further in strengthening my personal relationships, in enriching the relationships in my community, and in giving liberally of my talents and creativity. These are high virtues.

This is not to say that this is going to be easy. As Aristotle has said, the cultivation of virtue will most likely require a great amount of effort, work, and pain. We all seem so concerned about death that many of us fear expending effort and prefer, rather, to live a life of simple comforts or low aspirations—only to find ourselves, at age fifty or so, in mid-life crisis. We must realize, however, that death is nothing to fear. In a famous story of the Buddha, one of the Buddha's followers told the Buddha that he had not provided his followers with any knowledge of existence after death and that he wanted very much to know what happened to people after they die. In his wisdom, the Buddha replied to the man that the desire for such an inquiry proved that he was not free of suffering. The question concerning existence after death, in other words, demonstrated that the man still retained fear, anxiety, and unconscious or uncontrolled desires concerning death. Our souls might not be immortal, but our deeds are. Whatever we do makes an unalterable ripple in the flow of time and, thus, we live on in this way. This gives us all the more reason, therefore, for making certain, today, that those choices we make are the right ones. Most of us have no idea how much power is contained in the mere phenomenon of our own choice.

Some might offer up the comment that my account of the good is vacuous or uninteresting, for "all it gives us is tendential necessity." The critique here, in other words, is that the theory is too dry, cold, or formalistic. As Railton explains concerning the role of logic in ethics:

It could be said that I logically ought not to believe both a proposition p and a proposition that implies not- p . However, it may not be the case that every rational agent will have an instrumental reason to purge all logical contradictions from his thought. It would require

vast amounts of cognition for anyone to test all of his existing beliefs for consistency, and to insure that every newly acquired belief preserves it. Suppose someone to be so fortunate that the only contradictions among his beliefs lie deep in the much-sedimented swamp of factual trivia. Perhaps his memories of two past acquaintances have become confused in such a way that somewhere in the muck there are separate beliefs which, taken together, attribute to one individual logically incompatible properties. Until such a contradiction rears its head in practice, he may have no more reason to lay down his present concerns and wade in after it than he has to leave his home in suburban New Jersey to hunt alligators in the Okefenokee on the off chance that he might one day find himself stranded and unarmed in the backwaters of southeast Georgia. What an individual rationally ought to do thus may differ from what logic requires of him.⁴⁴

Many philosophers have thus criticized the view that *mere* logical consistency makes a decision ethical or right. In some sense, they are quite correct to make this critique. They have, however, missed the point. Logical consistency is important but not *primary* as a measurement of ethical soundness. That which is primary is, rather, that which is supremely valuable, the *summum bonum*. It is this *summum bonum* with which we must stay consistent, not, as Railton and others seem to suggest, logic itself. Thus, consistency is not merely about formality; there is rich content, rather, which consistency regulates. This relationship is akin to that in descriptive theory, for instance, regarding the mind-body problem. If one makes an argument concerning the immateriality of the mind, then all of the subsequent metaphysical propositions one sets forth must also be consistent with this immaterialist view of the mind. We do not, however, conclude from this that a reification or fetishization of consistency is at hand; we instead realize that there is rich content *behind* this consistency which the law of non-contradiction merely regulates. Furthermore, logical consistency is, as follows from the concept of best logical judgment, a constituent of the *telos* of every person. We should, therefore, indeed purge our minds of all contradictions. This is a necessity for a free and coherent mind. This view of normativity might, then, help to explain the kind of motivational internalism that Nagel adopts. Perhaps true virtue can motivate; indeed, I think it generally does. This thesis on motivational internalism or motivational consistency is an interesting and potentially very important one, and I stand as of yet undecided on the answer to the question of its significance to ethical thought.

The concept of teleologically comparative tendential necessity of functions, then, is exceedingly profound. What this concept entails is that there is a sound basis for our judgments, that we, as rational, conscious beings, can make choices that are *the right ones*, that are *the best ones*, and which are conducive to the highest, or best, possible purposes. We have, in other words, real, justifiable, and meaningful reasons for our actions, and we have objective reason to state that certain things in the world *must* be the way they are. The language of value is the language of rationality, and vice versa. Teleologically comparative tendential necessity, we must realize, is merely a measuring stick for evaluating value-

⁴⁴ Peter Albert Railton, "Moral Realism," *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Ed. Peter Albert Railton. (Cambridge: Cambridge University Press, 2003): 30.

theoretical phenomena, just as grams are a measure of things that actually have mass in the physical world. Thus, when we consider the multiplicity of things in the world to which comparative tendential necessity applies, we find that this concept is far from empty—it is, indeed, rich with value.

6.3.2.1 Preliminary Implications for Political Theory

As Macintyre and others have suggested, politics and society should be inseparable. What I think Macintyre means by this, and certainly what I mean by this, is that political decision-making should be localized, community-focused, and community-run. The principle of liberty derived from the *summum bonum* demands, in some contexts, collective, democratic decision-making and, in others, a system of private, individual decision-making, ultimately culminating in spontaneous cooperation. This is not the disorder and anarchy depicted by state maximalists. This, rather, is a system at once of real democracy, organized layer by layer according to region, locality, and voluntary participation, and also of real *laissez-faire*. It does not follow, however, that this must be a *state* system. Rather, what I look to in the future is a democratic society and political system in which representatives of the people are selected at will by whichever person or group of people they wish to represent them: no term limits, instantaneous voting, and instantaneous de-voting. Many people might, however, wish to eschew representation altogether and directly contribute to decisions within the community and within the larger political, inter-regional community. Others might seek alternatively to enjoin in interaction with others merely for purposes of trade or other personal satisfaction. Especially considering the state of the world as it is today, I believe we must be open to the serious consideration of such social arrangements.

There has been too much bickering between libertarian capitalists and libertarian socialists on the issue of which social system is better, but both sides, to their detriment, do not realize that they are indeed consistent with each other; the explanation for this seemingly unlikely state of affairs, however, must be held off until perhaps some other time. Suffice it to say for now that, in one community or region, workers' collectives and cooperatives might contribute to the greatest degree of liberation of the individuals who comprise it and that, in another community or region, individually owned, private enterprises might prove more beneficial. Indeed, I believe there is much to be said for the "democracy of the market." In addition, a community or region where there is a mixture of these two systems might also be most beneficial and, indeed, is probably inevitable. We must, however, be open to local social, economic, and technological experimentation, and this has already been happening to a small extent both in modern times and throughout history. Such libertarian-socialist institutions as Mondragon in Spain, John Lewis and Co-operative Group in Britain, the institution of participatory budgeting in Porto Alegre, and Christiania in Denmark, as well as such historical libertarian-socialist organizational arrangements as those in 1930s Barcelona, the quasi-autonomous organization of Harbin during the Chinese Revolution, and the

Free Territory during the Bolshevik Revolution, are prominent examples.⁴⁵ Unfortunately, modern examples of truly free-market systems, especially free of central banking, are in short supply, although there have been non-industrial systems, such as the relatively peaceful and equitable trade routes in ancient and medieval West Africa and the sociopolitical structures of ancient Ireland and Iceland, as well as the individualist social structures of various other ethnic groups, which seem to have approximated such systems.⁴⁶ This dearth of free-market examples, I believe, is mostly due to the massive government intervention in the economy (via the legal limitation of tort jurisdiction over polluters, via favoritism through the issuance of corporate charters, as well as through tariffs, war, and imperialism, via the establishment of central banking and fiat currencies, via copyright and patent laws, via the industrial-corporate control of legislative, judicial, and executive bodies, as well as other interventions) that was ushered in during the industrial revolutions, as well as during the period of neo-imperialist globalization in which we now live. From the United States, to Britain, to France, and even to such places as Egypt and Japan, industrialization was indeed high-jacked and perverted by state power. Of course, all of the aforementioned libertarian-socialist examples are instances of truly free-market systems, but sadly neither the right- nor the left-libertarians are keen on hearing out such an idea. For right-libertarians, all socialism is inherently oppressive, and for left-libertarians, anything with the words “free market” or “capitalism” is inherently evil.⁴⁷ Although the refutation of both of these theses would require an entire work on political theory, suffice it to say that I believe the mainstream right-libertarian and left-libertarian views to be fallacious and that it is really all forms of *state* socialism and *state* capitalism that are inherently oppressive. Left-libertarian Diego Abad de Santillan seems to have acknowledged the falsehood of the left-right dichotomy with regards to libertarianism. “In each locality,” he states “the degree of communism, collectivism or mutualism will depend on the conditions prevailing.

⁴⁵ There are certainly further examples, but these are a mere few. Other examples include but are not limited to: the democratically owned and managed Renmark Hotel in Australia (in operation since 1897), the long-standing cooperative tradition in Italy’s Emilia-Romagna region, Arden, Delaware and Jiaozhou Bay in Qingdao City and their historically Georgist-voluntarist models of social organization, the structure and composition of the workers’ councils in 1918 Germany, the co-housing tradition in Denmark, the long-standing Black Bear Ranch commune in the United States, the worker- and community-controlled enterprise systems in the Cleveland area studied by Gar Alperovitz, the rapidly expanding and highly effective violent-crime-prevention organization CeaseFire, and even the famous International in the nineteenth century. Of course, we should not blindly laud these organizations without any sort of criticism. Instead, we should undertake to understand the social factors and theoretical models at work within such systems so that we may avoid potential pitfalls and base the future organization of similar institutions on still sounder knowledge of human behavior.

⁴⁶ Bruce Benson, *The Enterprise of Law: Justice Without the State*. (Oakland: The Independent Institute, 2011); David Friedman, *The Machinery of Freedom: Guide to a Radical Capitalism*. (La Salle: Open Court Publishing, 1989); Murray Rothbard, *For a New Liberty: The Libertarian Manifesto*. (New York: Collier Macmillan Publishers, 2002).

⁴⁷ It is worth noting, however, that, in most instances, right-libertarians are extremely tolerant of the formation of socialist and communist organizational structures on a voluntary basis.

Why dictate rules? We who make freedom our banner, cannot deny it in economy. Therefore there must be free experimentation, free show of initiative and suggestions, as well as the freedom of organization.”⁴⁸ We must acknowledge, therefore, the real basis for the destruction of this dichotomy and, therefore, for the unification of the libertarian idea. The question that remains is: “What would such libertarian societies look like, and, if ethically required, how would we get there?” The even more important question is “Knowing what we do about the current abysmal realities of the world, and acknowledging the realities of the system as it currently exists, what must we do at this very instant to address them?” That is for another work, aided by intensive empirical and philosophical analysis, to answer.

My theory of the good is not a political theory, but, as is any ethical theory, it is a jumping-off point for political theory. As humans, we are imperfect, and, although we probably cannot create a utopia, we should still strive to create one; indeed, evidence abounds in the world today of a need for radical social and economic change. Let us not, however, be fooled by those who consider themselves adherents to *realpolitik*. Real theories of the good yield real politics, while legal positivism yields nothing but an illusory mask for elitism and prejudice. We can create a virtuous society, but we cannot *act* before we *think*. Let us not, however, be fooled. For every new theory or discourse, there must be action. We must not be afraid of having ideals or be afraid of being mocked for having ideals. Those who are relativist or relativist-leaning will tell us that such ideals are merely exuberant visions and idle dreams. But they forget that great acts throughout history have necessitated great ideals without which many of those acts would not have been accomplished. The revolutionary overthrow of tyrannical regimes and the transition to a more autonomous existence in the Middle East, the resistance of the people of Hong Kong (whose vociferous love of liberty is quite explicit), Tibet, and others to the oppression of the Chinese government, the Occupy movements around the world (which are comprised of both left- and right-libertarian strains) and the current rumblings for change in the U.S., Europe, Asia, Africa, and elsewhere, the various anti-imperialist and anti-slavery movements both past and contemporary, and the many other struggles for civil rights throughout history necessitate ideals, virtues. We often forget that ideas and institutions that we take for granted today were, in historically recent periods, considered absurd, impractical, and utopian. That people of color can exist free alongside free white people was once widely considered ridiculous and dangerous, and that women can do the same as regards men was considered inane. That nations could unite in an international forum, however flawed, for settling their disputes and discussing common interests was merely the result of exuberant conjurings on the part of philosophers; that people across the globe could communicate instantaneously and form cross-cultural and cross-national solidarity was folly. Indeed, the medicines and technologies we now possess were too far from the realm of the imagination even to be dismissed as romantic fancy. We must remember, too, that but two centuries ago the idea of

⁴⁸ Robert Graham, *Anarchism: A Documentary History of Libertarian Ideas, Volume One*. (Castroville: Black Rose Books, 2004): 475.

democratic republicanism was laughed at and ridiculed by those in power as a fantastic, idealistic utopianism. While this idea was and still is undoubtedly better than the previous alternatives, it is abundantly and painfully clear, despite the continued attitudes of those in denial, that democratic republicanism, whether state-capitalist or state-socialist, has failed.⁴⁹ Throughout the history of the advancement of democratic republicanism, there have always been those who have longed to bring to fruition, and who on occasion have successfully brought to fruition, a vision of society which advances the principle of liberty still further. It is today that we must all take up this mantle.

Every action must have, as its goal, a justice based upon universal equality of liberty, a real, universal discourse, and an eternal eye toward the good. It is probably not the case that everyone, or even perhaps the majority of people, will become virtuous or excellent; but there are many now who, oppressed, would become excellent if released from their oppression. Most people, I believe, naturally want to do good; all they need is the tools to do so. Once we become responsible to each other, once we light the fire of universal discourse and truly care about each other, the shackles will fall, the sun of enlightenment will rise, and we will show relativists, nihilists, and the realpolitikers that true liberation is possible and that a virtuous and free society is not simply an idle dream.

⁴⁹ Let us not, however, give fuel to the fire of tyrants; while democratic republicanism is deeply problematic, it is certainly a step up from totalitarianism, dictatorship, or anocracy. My comment, therefore, does not give an excuse to elites wishing not to relinquish power in favor of a democratic-republican alternative and to the subversion of the obvious wish of subjugated peoples around the world to engage in such a transition. Still here, we must be careful. I do not pose a claim to some dialectical or otherwise developmental theory of history, as Marx and others have; my position is far less presumptuous. Whether statist reformism or outright revolution for the establishment of democratic republicanism, or an outright libertarian revolution of peaceful resistance and disobedience to State-created law, it is the people themselves who must decide which route at which moment is most viable. I am confident, however, that they too will ultimately realize both the tyranny and the unsustainability of democratic republicanism.

Chapter 7

Conclusion

7.1 The Eternal Nature of the *Summum Bonum*

There are theorists who have reasoned that, if there are laws of physics, laws of mathematics, and laws of logic, it is not out of the question to posit laws of morality or ethics. Indeed, natural-law theorists took and still take this position with regards to ethics. The positive-law tradition, of course, overshadowed the natural-law tradition in Europe in the nineteenth century, with its supposed pragmatism and practicality. This pragmatism and practicality, of course, along with other so-called pragmatisms, such as Bismarckian *realpolitik*, exposed themselves throughout the century and beyond as mere prejudices conducive only to horror and destruction. This is because legal positivism is actually a euphemism for relativism or subjectivism, both of which, when injected into a culture, must inevitably end in such horror and destruction. Positivism and relativism, we must finally admit, are neither pragmatic nor practical, and, although theorizing has been criticized for its impracticality, we must see that it is actually far more practical than any legal positivism or ethical relativism. Maybe, however, the positivists were right about one thing: theorizing, alone, does not achieve results. Let us not, then, drop theorizing altogether, which was the terrible mistake of positivism; let us, rather, combine theory with action. “Why?,” we should ask, “are relativism and subjectivism so dangerous?” The answer is simple: their number-one danger is in their promotion of complacency. If nobody’s ethical claims are right or wrong, then there can be no legitimate criticism of any particular act. If Stalin thinks it is okay to imprison, torture, and execute supposed Trotskyites, or if someone of a particular group wants to subject someone else to genital or other bodily mutilation against her will, or if a government wants to commit car bombings and other terrorist attacks against civilians, or if Lot wants to submit his daughters to a mob to be raped, or if someone wants to bind a woman’s feet, or if someone wants to enslave someone else because she or he is black, or if someone’s religious beliefs impel her to despise all LGBTQ people, or if someone wants simply to tell a series of lies to cheat someone else out of the money he or she needs to live, then there is no legitimate criticism we can make of any of these beliefs or actions. We get stuck in

the social rut of making excuses for the evils in people's cultural practices, religions, or overall views, which include, but are not limited to, such excuses as "But it is their culture," "It is their religion," or "It is the way he wants to live his life." This, of course, does not give us an excuse, for instance, to use force or violence in order to curtail such evil. In previous chapters, the reader has gotten a glimpse at those rare times when such force is justifiable and even obligatory. In most cases, however, it is discourse or mediation—not force—which is necessary. Especially as Habermas's theory has demonstrated, it would be wrong to engage in violence or coercion in order to enforce our opinions upon others. In opposition to the relativist, we have, indeed, seen that there is objective goodness and even objective bestness and, furthermore, that the *summum bonum* is eternal. Philosophers often implicitly deride such a view. Once again, as Railton states: "The notions of good or bad have a place in the scheme of things only in virtue of facts about what matters or could matter to beings for whom it is possible that something matter. A being for whom something can matter is a being with a point of view, a subjectivity. In a universe without subjectivity, there is no value either."¹ This view, however, oversimplifies the issue. What do I mean when I say that the *summum bonum* is eternal? As mathematical laws and logical laws hold eternally, whether there are agents in the universe or not, the laws of morality or ethics hold eternally. Just as the relationship $((P \rightarrow Q) \rightarrow (\neg P \vee Q))$ holds whether or not there are agents, so does the relationship "Rationality is the most tendentially necessary for the teleologically comparative tendential necessity of functions" or "Rationality is the best for betterness" hold, whether or not there are agents. Few would make the bold claim, for instance, that, if all sapient beings in the universe went extinct, the earth would suddenly lose its circumference or that the trigonometric properties of radio or infrared waves would disappear. We do not, then, invent this value relationship; instead, we come to realize it. To compare something as better or worse than something else is to make a claim, not concerning some arbitrary *telos*, but concerning the essence and *telos* of betterness. Thus, if such relationships as $((P \rightarrow Q) \rightarrow (\neg P \vee Q))$ are the case, have always been the case, and always will be the case, then not only is rationality the *summum bonum*, but it always has been the *summum bonum*, and it always will be. Without this truth, it would be incoherent to state that, if there were no rational beings in the universe, the universe would be an abysmal and tragic place. This, indeed, is true.

7.2 Enlightenment and Liberation

The claim will inevitably arise, once again, that rationality's status as the *summum bonum* is elitist, and I believe it is, once more, important to tackle this concern. It is not the case, for instance, that everyone can achieve all of the goods necessary for best rationality. This, we must admit, is true. There are, and have been, oppressed

¹ Railton, "Facts and Values," 56.

people throughout the world who are in their condition because of discrimination, ignorance on the part of others, some form of elitism, or all three. This, however, does not change the fact that rationality is the *summum bonum*, and it certainly does not mark oppressed people as inferior. On the contrary, rationality's status as the *summum bonum* becomes, for us, a directive—a directive to contribute to the excellence of oppressed people by aiding them. Let us be emphatically clear, as well, that this is not some paternalistic directive, a directive which has, throughout history, more often resulted in a greater lack of freedom for its patients. Unlike with the imperialistic and paternalistic approaches of the past, I have pointed out that it is not simply better logical rationality that makes someone more excellent; it is, rather, logical rationality *and* experiential rationality, and only oppressed people, themselves, and all individual people in general, know their own experiences. We cannot tell people what they experience or how they experience. We must, rather, engage in discourse with oppressed people and try to understand their lives. Only if we listen can we understand, and only the individual herself or himself can find his or her rational *telos*. We can aid others in finding their rational *telos*, but, in the end, it is only they who can enlighten themselves.

We must not neglect, of course, the oppression of non-human animals and even the destruction and abuse of other forms of life. In the film “La Strada,” by Federico Fellini, the character, Il Matto, tells Gelsomina that everything has a purpose, “even this pebble. . . even the stars.” He was right, and our previous analyses have shown this to be the logical conclusion of a consideration of formal and final causality. We can help all living things achieve their *teloi* by living virtuously ourselves.

Rational beings, however, need more than survival for the fulfillment of their *teloi*.² They need certain virtues that are specifically rational ones: self-awareness, self-knowledge, self-discipline, resiliency, creativity, compassion, and the ability to engage in discourse (broadly construed), among others. These virtues, we see, aid in yielding better individual autonomy and freedom, better conceptual synthesis, better conceptual abstraction, better judgment, and a better life overall. If these faculties of rationality are fulfilled in an individual to that individual's greatest capacity, then that individual is living the best life that she or he can possibly live. The excellence of rationality in an individual life, therefore, is both necessary and sufficient for what has traditionally been called “the good life.”

Such philosophers as Macintyre, Habermas, Apel, and Sen have stressed the importance of discourse in ethics, and they are right to do so. The importance of free and open discourse cannot be stressed enough. Indeed, it is due in great part to the lack of such discourse in a society that prejudices of various kinds exist. We live not in an open society but in an isolated society, enforced and perpetuated by what Habermas, Foucault, and others have exposed as institutionalized structures of power and domination. Because of such hypostasized powers and power structures as exist

² I do not here mean to imply that all non-rational living beings have only survival as their *telos*. In order to determine their respective *teloi*, of course, we would have to engage in serious inquiry into their respective essences.

in the political system, the legal system, the financial system, the labor system, and in the professions, discursive and cooperative incentives have become marginalized. It is perhaps because of such power structures that, upon a superficial analysis of society, our interests seem so disparate. But, as we have seen, my interests cannot be fulfilled without yours also being fulfilled; it is fallacious for me to claim that I have become excellent if, in every step I take, I have not advanced further in strengthening my personal relationships, in enriching the relationships in my community, and in giving liberally of my talents and creativity. These are high virtues.

This is not to say that this is going to be easy. As Aristotle has said, the cultivation of virtue will most likely require a great amount of effort, work, and pain. We all seem so concerned about death that many of us fear expending effort and prefer, rather, to live a life of simple comforts or low aspirations—only to find ourselves, at age 50 or so, in mid-life crisis. We must realize, however, that death is nothing to fear. In a famous story of the Buddha, one of the Buddha's followers told the Buddha that he had not provided his followers with any knowledge of existence after death and that he wanted very much to know what happened to people after they die. In his wisdom, the Buddha replied to the man that the desire for such an inquiry proved that he was not free of suffering. The question concerning existence after death, in other words, demonstrated that the man still retained fear, anxiety, and unconscious or uncontrolled desires concerning death. Our souls might not be immortal, but our deeds are. Whatever we do makes an unalterable ripple in the flow of time and, thus, we live on in this way. This gives us all the more reason, therefore, for making certain, today, that those choices we make are the right ones. Most of us have no idea how much power is contained in the mere phenomenon of our own choice.

Some might offer up the comment that my account of the good is vacuous or uninteresting, for “all it gives us is tendential necessity.” The critique here, in other words, is that the theory is too dry, cold, or formalistic. As Railton explains concerning the role of logic in ethics:

It could be said that I logically ought not to believe both a proposition p and a proposition that implies not- p . However, it may not be the case that every rational agent will have an instrumental reason to purge all logical contradictions from his thought. It would require vast amounts of cognition for anyone to test all of his existing beliefs for consistency, and to insure that every newly acquired belief preserves it. Suppose someone to be so fortunate that the only contradictions among his beliefs lie deep in the much-sedimented swamp of factual trivia. Perhaps his memories of two past acquaintances have become confused in such a way that somewhere in the muck there are separate beliefs which, taken together, attribute to one individual logically incompatible properties. Until such a contradiction rears its head in practice, he may have no more reason to lay down his present concerns and wade in after it than he has to leave his home in suburban New Jersey to hunt alligators in the Okefenokee on the off chance that he might one day find himself stranded and unarmed in the backwaters of southeast Georgia. What an individual rationally ought to do thus may differ from what logic requires of him.³

³Peter Albert Railton, “Moral Realism,” *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Ed. Peter Albert Railton. (Cambridge: Cambridge University Press, 2003): 30.

Many philosophers have thus criticized the view that *mere* logical consistency makes a decision ethical or right. In some sense, they are quite correct to make this critique. They have, however, missed the point. Logical consistency is important but not *primary* as a measurement of ethical soundness. That which is primary is, rather, that which is supremely valuable, the *summum bonum*. It is this *summum bonum* with which we must stay consistent, not, as Railton and others seem to suggest, logic itself. Thus, consistency is not merely about formality; there is rich content, rather, which consistency regulates. This relationship is akin to that in descriptive theory, for instance, regarding the mind-body problem. If one makes an argument concerning the immateriality of the mind, then all of the subsequent metaphysical propositions one sets forth must also be consistent with this immaterialist view of the mind. We do not, however, conclude from this that a reification or fetishization of consistency is at hand; we instead realize that there is rich content *behind* this consistency which the law of non-contradiction merely regulates. Furthermore, logical consistency is, as follows from the concept of best logical judgment, a constituent of the *telos* of every person. We should, therefore, indeed purge our minds of all contradictions. This is a necessity for a free and coherent mind. This view of normativity might, then, help to explain the kind of motivational internalism that Nagel adopts. Perhaps true virtue can motivate; indeed, I think it generally does. This thesis on motivational internalism or motivational consistency is an interesting and potentially very important one, and I stand as of yet undecided on the answer to the question of its significance to ethical thought.

The concept of teleologically comparative tendential necessity of functions, then, is exceedingly profound. What this concept entails is that there is a sound basis for our judgments, that we, as rational, conscious beings, can make choices that are *the right ones*, that are *the best ones*, and which are conducive to the highest, or best, possible purposes. We have, in other words, real, justifiable, and meaningful reasons for our actions, and we have objective reason to state that certain things in the world *must* be the way they are. The language of value is the language of rationality, and vice versa. Teleologically comparative tendential necessity, we must realize, is merely a measuring stick for evaluating value-theoretical phenomena, just as grams are a measure of things that actually have mass in the physical world. Thus, when we consider the multiplicity of things in the world to which comparative tendential necessity applies, we find that this concept is far from empty—it is, indeed, rich with value.

7.3 Preliminary Implications for Political Theory

As Macintyre and others have suggested, politics and society should be inseparable. What I think Macintyre means by this, and certainly what I mean by this, is that political decision-making should be localized, community-focused, and community-run. The principle of liberty derived from the *summum bonum* demands, in some contexts, collective, democratic decision-making and, in others, a system of private, individual decision-making, ultimately culminating in

spontaneous cooperation. This is not the disorder and anarchy depicted by state maximalists. This, rather, is a system at once of real democracy, organized layer by layer according to region, locality, and voluntary participation, and also of real *laissez-faire*. It does not follow, however, that this must be a *state* system. Rather, what I look to in the future is a democratic society and political system in which representatives of the people are selected at will by whichever person or group of people they wish to represent them: no term limits, instantaneous voting, and instantaneous de-voting. Many people might, however, wish to eschew representation altogether and directly contribute to decisions within the community and within the larger political, inter-regional community. Others might seek alternatively to enjoin in interaction with others merely for purposes of trade or other personal satisfaction. Especially considering the state of the world as it is today, I believe we must be open to the serious consideration of such social arrangements.

There has been too much bickering between libertarian capitalists and libertarian socialists on the issue of which social system is better, but both sides, to their detriment, do not realize that they are indeed consistent with each other; the explanation for this seemingly unlikely state of affairs, however, must be held off until perhaps some other time. Suffice it to say for now that, in one community or region, workers' collectives and cooperatives might contribute to the greatest degree of liberation of the individuals who comprise it and that, in another community or region, individually owned, private enterprises might prove more beneficial. Indeed, I believe there is much to be said for the "democracy of the market." In addition, a community or region where there is a mixture of these two systems might also be most beneficial and, indeed, is probably inevitable. We must, however, be open to local social, economic, and technological experimentation, and this has already been happening to a small extent both in modern times and throughout history. Such libertarian-socialist institutions as Mondragon in Spain, John Lewis and Co-operative Group in Britain, the institution of participatory budgeting in Porto Alegre, and Christiania in Denmark, as well as such historical libertarian-socialist organizational arrangements as those in 1930s Barcelona, the quasi-autonomous organization of Harbin during the Chinese Revolution, and the Free Territory during the Bolshevik Revolution, are prominent examples.⁴ Unfortunately, modern examples of truly free-market systems, especially free of central banking, are in short supply, although there have been non-industrial

⁴ There are certainly further examples, but these are a mere few. Other examples include but are not limited to: the democratically owned and managed Renmark Hotel in Australia (in operation since 1897), the long-standing cooperative tradition in Italy's Emilia-Romagna region, Arden, Delaware and Jiaozhou Bay in Qingdao City and their historically Georgist-voluntarist models of social organization, the structure and composition of the workers' councils in 1918 Germany, the co-housing tradition in Denmark, the long-standing Black Bear Ranch commune in the United States, the worker- and community-controlled enterprise systems in the Cleveland area studied by Gar Alperovitz, the rapidly expanding and highly effective violent-crime-prevention organization CeaseFire, and even the famous International in the nineteenth century. Of course, we should not blindly laud these organizations without any sort of criticism. Instead, we should undertake to understand the social factors and theoretical models at work within such systems so that we may avoid potential pitfalls and base the future organization of similar institutions on still sounder knowledge of human behavior.

systems, such as the relatively peaceful and equitable trade routes in ancient and medieval West Africa and the sociopolitical structures of ancient Ireland and Iceland, as well as the individualist social structures of various other ethnic groups, which seem to have approximated such systems.⁵ This dearth of free-market examples, I believe, is mostly due to the massive government intervention in the economy (via the legal limitation of tort jurisdiction over polluters, via favoritism through the issuance of corporate charters, as well as through tariffs, war, and imperialism, via the establishment of central banking and fiat currencies, via copyright and patent laws, via the industrial-corporate control of legislative, judicial, and executive bodies, as well as other interventions) that was ushered in during the industrial revolutions, as well as during the period of neo-imperialist globalization in which we now live. From the United States, to Britain, to France, and even to such places as Egypt and Japan, industrialization was indeed high-jacked and perverted by state power. Of course, all of the aforementioned libertarian-socialist examples are instances of truly free-market systems, but sadly neither the right- nor the left-libertarians are keen on hearing out such an idea. For right-libertarians, all socialism is inherently oppressive, and for left-libertarians, anything with the words “free market” or “capitalism” is inherently evil.⁶ Although the refutation of both of these theses would require an entire work on political theory, suffice it to say that I believe the mainstream right-libertarian and left-libertarian views to be fallacious and that it is really all forms of *state* socialism and *state* capitalism that are inherently oppressive. Left-libertarian Diego Abad de Santillan seems to have acknowledged the falsehood of the left-right dichotomy with regards to libertarianism. “In each locality,” he states “the degree of communism, collectivism or mutualism will depend on the conditions prevailing. Why dictate rules? We who make freedom our banner, cannot deny it in economy. Therefore there must be free experimentation, free show of initiative and suggestions, as well as the freedom of organization.”⁷ We must acknowledge, therefore, the real basis for the destruction of this dichotomy and, therefore, for the unification of the libertarian idea. The question that remains is: “What would such libertarian societies look like, and, if ethically required, how would we get there?” The even more important question is “Knowing what we do about the current abysmal realities of the world, and acknowledging the realities of the system as it currently exists, what must we do at this very instant to address them?” That is for another work, aided by intensive empirical and philosophical analysis, to answer.

My theory of the good is not a political theory, but, as is any ethical theory, it is a jumping-off point for political theory. As humans, we are imperfect, and, although we

⁵ Bruce Benson, *The Enterprise of Law: Justice Without the State*. (Oakland: The Independent Institute, 2011); David Friedman, *The Machinery of Freedom: Guide to a Radical Capitalism*. (La Salle: Open Court Publishing, 1989); Murray Rothbard, *For a New Liberty: The Libertarian Manifesto*. (New York: Collier Macmillan Publishers, 2002).

⁶ It is worth noting, however, that, in most instances, right-libertarians are extremely tolerant of the formation of socialist and communist organizational structures on a voluntary basis.

⁷ Robert Graham, *Anarchism: A Documentary History of Libertarian Ideas, Volume One*. (Castroville: Black Rose Books, 2004): 475.

probably cannot create a utopia, we should still strive to create one; indeed, evidence abounds in the world today of a need for radical social and economic change. Let us not, however, be fooled by those who consider themselves adherents to *realpolitik*. Real theories of the good yield real politics, while legal positivism yields nothing but an illusory mask for elitism and prejudice. We can create a virtuous society, but we cannot *act* before we *think*. Let us not, however, be fooled. For every new theory or discourse, there must be action. We must not be afraid of having ideals or be afraid of being mocked for having ideals. Those who are relativist or relativist-leaning will tell us that such ideals are merely exuberant visions and idle dreams. But they forget that great acts throughout history have necessitated great ideals without which many of those acts would not have been accomplished. The revolutionary overthrow of tyrannical regimes and the transition to a more autonomous existence in the Middle East, the resistance of the people of Hong Kong (whose vociferous love of liberty is quite explicit), Tibet, and others to the oppression of the Chinese government, the Occupy movements around the world (which are comprised of both left- and right-libertarian strains) and the current rumblings for change in the U.S., Europe, Asia, Africa, and elsewhere, the various anti-imperialist and anti-slavery movements both past and contemporary, and the many other struggles for civil rights throughout history necessitate ideals, virtues. We often forget that ideas and institutions that we take for granted today were, in historically recent periods, considered absurd, impractical, and utopian. That people of color can exist free alongside free white people was once widely considered ridiculous and dangerous, and that women can do the same as regards men was considered inane. That nations could unite in an international forum, however flawed, for settling their disputes and discussing common interests was merely the result of exuberant conjurings on the part of philosophers; that people across the globe could communicate instantaneously and form cross-cultural and cross-national solidarity was folly. Indeed, the medicines and technologies we now possess were too far from the realm of the imagination even to be dismissed as romantic fancy. We must remember, too, that but two centuries ago the idea of democratic republicanism was laughed at and ridiculed by those in power as a fantastic, idealistic utopianism. While this idea was and still is undoubtedly better than the previous alternatives, it is abundantly and painfully clear, despite the continued attitudes of those in denial, that democratic republicanism, whether state-capitalist or state-socialist, has failed.⁸ Throughout the history of the advancement of

⁸ Let us not, however, give fuel to the fire of tyrants; while democratic republicanism is deeply problematic, it is certainly a step up from totalitarianism, dictatorship, or anocracy. My comment, therefore, does not give an excuse to elites wishing not to relinquish power in favor of a democratic-republican alternative and to the subversion of the obvious wish of subjugated peoples around the world to engage in such a transition. Still here, we must be careful. I do not pose a claim to some dialectical or otherwise developmental theory of history, as Marx and others have; my position is far less presumptuous. Whether statist reformism or outright revolution for the establishment of democratic republicanism, or an outright libertarian revolution of peaceful resistance and disobedience to State-created law, it is the people themselves who must decide which route at which moment is most viable. I am confident, however, that they too will ultimately realize both the tyranny and the unsustainability of democratic republicanism.

democratic republicanism, there have always been those who have longed to bring to fruition, and who on occasion have successfully brought to fruition, a vision of society which advances the principle of liberty still further. It is today that we must all take up this mantle.

Every action must have, as its goal, a justice based upon universal equality of liberty, a real, universal discourse, and an eternal eye toward the good. It is probably not the case that everyone, or even perhaps the majority of people, will become virtuous or excellent; but there are many now who, oppressed, would become excellent if released from their oppression. Most people, I believe, naturally want to do good; all they need is the tools to do so. Once we become responsible to each other, once we light the fire of universal discourse and truly care about each other, the shackles will fall, the sun of enlightenment will rise, and we will show relativists, nihilists, and the realpolitikers that true liberation is possible and that a virtuous and free society is not simply an idle dream.

Bibliography

- Alexander, Richard D. 1993. Biological considerations in the analysis of morality. In *Evolutionary ethics*, ed. Matthew H. Nitecki and Doris V. Nitecki, 163–196. Albany: State University of New York Press.
- Allison, H. 1990. *Kant's theory of freedom*. Cambridge: Cambridge University Press.
- Amiot, C.E., et al. 2012. Can intergroup behaviors be emitted out of self-determined reasons? Testing the role of group norms and behavioral congruence in the internalization of discrimination and parity behaviors. *Personality and Social Psychology Bulletin* 38: 63–76.
- Anscombe, G.E.M. 1997. Modern moral philosophy. In *Virtue ethics*, ed. Roger Crisp and Michael Slote, 26–45. Oxford: Oxford University Press.
- Anscombe, G.E.M. 2000. *Intention*. Cambridge: Harvard University Press.
- Anscombe, G.E.M. 2005. *Human life, action and ethics: Essays*. Charlottesville: Imprint Academic.
- Apel, Karl-Otto. 1996a. The question of grounding: Philosophy and transcendental pragmatics of language. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 68–102. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 1996b. Types of rationality today: The continuum of reason between science and ethics. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 137–173. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 1996c. The *Conditio Humana* as an ethical problem. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 174–191. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 1996d. Limits of discourse ethics? An attempt at a provisional assessment. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 192–218. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 1996e. The challenge of a totalizing critique of reason and the program of a philosophical theory of rationality types. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 250–274. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 1996f. The rationality of human communication: On the relationship between consensual, strategic, and systems rationality. In *Karl-Otto Apel: Selected essays, vol. 2: Ethics and the theory of rationality*, ed. Eduardo Mendieta, 316–339. Atlantic Highlands: Humanities Press.
- Apel, Karl-Otto. 2001. *La Réponse de L'Ethique de la Discussion au Défi Moral de la Situation Humaine Comme Telle et Spécialement Aujourd'hui*. Louvain: Editions de l'Institut Supérieur de Philosophie Louvain-la-Neuve.
- Aristotle. 2004. *Nicomachean ethics*, ed. Roger Crisp. Cambridge: Cambridge University Press.

- Arrhenius, G. 2005. Superiority in value. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 291–304. Dordrecht: Springer.
- Augustine. 1998. *Confessions*. Oxford: Oxford University Press.
- Bandura, Albert. 1985. *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River: Prentice Hall.
- Bandura, Albert. 1995. *Self-efficacy in changing societies*. Cambridge: Cambridge University Press.
- Bandura, Albert. 1997. *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.
- Baron, Marcia. 2011. Virtue ethics in relation to Kantian ethics: An opinionated overview and commentary. In *Perfecting virtue: New essays on Kantian ethics and virtue ethics*, ed. Lawrence Jost and Julian Wuerth, 8–38. Cambridge: Cambridge University Press.
- Beardsley, Monroe. 1965. Intrinsic value. *Philosophy and Phenomenological Research* 26(1): 1–17.
- Benhabib, Seyla. 1986. *Critique, norm, and utopia: A study of the foundations of critical theory*. New York: Columbia University Press.
- Benson, Bruce. 2011. *The enterprise of law: Justice without the state*. Oakland: The Independent Institute.
- Beyleveld, Deryck. 1991. *The dialectical necessity of morality: An analysis and defense of Alan Gewirth's argument to the principle of generic consistency*. Chicago: The University of Chicago Press.
- Beyleveld, Deryck. 1994. *Law as a moral judgment*. Sheffield: Sheffield Academic Press.
- Beyleveld, Deryck. 2000. Is embryo research and preimplantation genetic diagnosis ethical? *Forensic Science International* 113: 461–475.
- Beyleveld, Deryck. 2002. A reply to Marcus G. Singer on Gewirth, Beyleveld and dialectical necessity. *Ratio Juris* 15(4): 458–473.
- Beyleveld, Deryck. 2008. Human cognitive vulnerability and the moral status of the human embryo and foetus. In *The contingent nature of life*, ed. Marcus Duwell, Christoph Rehmann-Sutter, and Dietmar Mieth, 83–88. Dordrecht: Springer.
- Beyleveld, Deryck. 2012. Emerging technologies, extreme uncertainty, and the principle of rational precautionary reasoning. *Law, Innovation and Technology* 4(1): 35–65.
- Beyleveld, Deryck, and Roger Brownsword. 2001. *Human dignity in bioethics and biolaw*. Oxford: Oxford University Press.
- Beyleveld, Deryck, and Roger Brownsword. 2006. Principle, proceduralism, and precaution in a community of rights. *Ratio Juris* 19(2): 141–168.
- Beyleveld, Deryck, and Shaun Pattinson. 2000. Precautionary reasoning as a link to moral action. In *Medical ethics*, ed. Michael Boylan, 39–53. Upper Saddle River: Prentice Hall.
- Blanshard, Brand. 1961. *Reason and goodness*. Atlantic Highlands: Humanities Press.
- Blewett, Lynn A., et al. 2004. Monitoring the uninsured: A state policy perspective. *Journal of Health Politics, Policy and Law* 29(1): 107–145.
- Bodanszky, Eva, and Earl Conee. 1981. Isolating intrinsic value. *Analysis* 41(1): 51–53.
- Bond, Edward J. 1980a. Gewirth on reason and morality. *Metaphilosophy* 11: 36–53.
- Bond, Edward J. 1980b. Reply to Gewirth. *Metaphilosophy* 11: 70–75.
- Bostock, David. 2000. *Aristotle's ethics*. New York: Oxford University Press.
- Boylan, Michael. 1999. *Gewirth: Critical essays on action, rationality, and community*. Lanham: Rowman & Littlefield.
- Brink, David. 1989. *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press.
- Brook, Andrew. 1994. *Kant and the mind*. Cambridge: Cambridge University Press.
- Brown, Stephen A. 1999. An antidote to value-free conceptions. *Studies in Social and Political Thought* 1: 86–92.
- Bubner, Rüdiger. 1982. Habermas's concept of critical theory. In *Habermas: Critical debates*, ed. John B. Thompson and David Held. Cambridge: The MIT Press.

- Bundorf, M. Kate, and Mark V. Pauly. 2006. Is health insurance affordable for the uninsured? *Journal of Health Economics* 25: 650–673.
- Carlson, Erik. 1996. The intrinsic value of non-basic states of affairs. *Philosophical Studies* 85(1): 95–107.
- Carnap, Rudolph. 1950. Empiricism, semantics, ontology. *Revue Internationale de Philosophie* 4: 20–40.
- Chappell, Timothy. 2011. Intuition, system, and the ‘paradox’ of deontology. In *Perfecting virtue: New essays on Kantian ethics and virtue ethics*, ed. Lawrence Jost and Julian Wuerth, 271–289. Cambridge: Cambridge University Press.
- Chisholm, R.M. 2005a. Intrinsic value. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 1–10. Dordrecht: Springer.
- Chisholm, R.M. 2005b. Defining intrinsic value. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 15–16. Dordrecht: Springer.
- Chomsky, Noam. 1977. *On language*. New York: The New Press.
- Claassen, Rutger, and Marcus Düwell. 2012. The foundations of capability theory: Comparing Nussbaum and Gewirth. *Ethical Theory and Moral Practice* 16(3): 493–510.
- Danielsson, Sven. 1998. Harman’s equation and non-basic intrinsic value. *Uppsala Philosophical Studies* 47: 11–19.
- Darwell, Stephen, Allan Gibbard, and Peter Railton. 1997. *Moral discourse & practice: Some philosophical approaches*. Oxford: Oxford University Press.
- Davis, Michael. 1991. Gewirth and the pain of contradiction. *Philosophical Forum* 22: 211–227.
- Donagan, Alan. 1977. *The theory of morality*. Chicago: The University of Chicago Press.
- Doris, John M. 2002. *Lack of character: Personality & moral behavior*. Cambridge: Cambridge University Press.
- Doviak, Daniel. 2011. A new form of agent-based virtue ethics. *Ethical Theory and Moral Practice*. 14: 259–272.
- Elzanowski, Andrzej. 1993. The moral career of vertebrate values. In *Evolutionary ethics*, ed. Matthew H. Nitecki and Doris V. Nitecki, 259–276. Albany: State University of New York Press.
- Engstrom, S. 1992. The concept of the highest good in Kant’s moral theory. *Philosophy and Phenomenological Research* 52(4): 747–780.
- Enoch, David. 2011. *Taking morality seriously: A defense of Robust realism*. Oxford: Oxford University Press.
- Epicurus. 2007. *Principal doctrines*, ed. Robert Drew Hicks. E-Text: <http://ebooks.adelaide.edu.au/e/epicurus/doctrines/>
- Feit, Neil. 2001. The structure of higher goods. *Southern Journal of Philosophy* 39(1): 47–57.
- Feldman, Fred. 1998. Hyperventilating about intrinsic value. *Journal of Ethics*. 2(4): 339–354.
- Feldman, Fred. 2000. Basic intrinsic value. *Philosophical Studies* 99(3): 319–346.
- Foot, Philippa. 1958–1959. Moral beliefs. *Proceedings of the Aristotelian Society* 59: 83–104. Reprinted in Philippa Foot, ed., *Theories of ethics*, 83–100.
- Foot, Philippa. 2001. *Natural goodness*. Oxford: Oxford University Press.
- Foot, Philippa. 2002a. Goodness and choice. In *Virtues and vices: And other essays in moral philosophy*, ed. Philippa Foot, 132–145. Oxford: Oxford University Press.
- Foot, Philippa. 2002b. Moral arguments. In *Virtues and vices: And other essays in moral philosophy*, ed. Philippa Foot, 96–109. Oxford: Oxford University Press.
- Forst, Rainer. 2012. *The right to justification: Elements of a constructivist theory of justice*. New York: Columbia University Press.
- Frankena, William. 1939. The naturalistic fallacy. *Mind* 48: 464–477.
- Friedman, David. 1989. *The machinery of freedom: Guide to a radical capitalism*. La Salle: Open Court Publishing.
- Gauthier, David. 1986. *Morals by agreement*. Oxford: Oxford University Press.
- Gazzaniga, Michael S. 2005. *The ethical brain*. New York: Dana Press.
- Geach, Peter. 1958. Imperative and deontic logic. *Analysis* 18(3): 49–56.

- Geach, Peter. 1967. Good and evil. In *Theories of ethics*, ed. Philippa Foot, 64–73. Oxford: Oxford University Press.
- Geertz, Clifford. 1973. *The interpretation of cultures*. New York: Basic Books.
- Gert, Bernard. 1998. *Morality: Its nature and justification*. Oxford: Oxford University Press.
- Gewirth, Alan. 1973. The is-ought problem resolved. *Proceedings and Addresses of the American Philosophical Association* 47: 34–61.
- Gewirth, Alan. 1978. *Reason and morality*. Chicago: The University of Chicago Press.
- Gewirth, Alan. 1983. *Human rights: Essays on justification and applications*. Chicago: University of Chicago Press.
- Gewirth, Alan. 1985. Rights and virtues. *The Review of Metaphysics* 38(4): 739–762.
- Gewirth, Alan. 1993. How ethical is evolutionary ethics? In *Evolutionary ethics*, ed. Matthew H. Nitecki and Doris V. Nitecki, 242–256. Albany: State University of New York Press.
- Gewirth, Alan. 1996. *The community of rights*. Chicago: The University of Chicago Press.
- Gewirth, Alan. 1998. *Self-fulfillment*. Princeton: Princeton University Press.
- Gibbard, Alan. 1992. *Wise choices, apt feelings: A theory of normative judgment*. Cambridge: Harvard University Press.
- Godwin, William. 1842. *Enquiry concerning political justice and its influence on morals and happiness*. London: J. Watson.
- Gómez-Lobo, Alfonso. 1989. The ergon inference. *Phronesis* 34(2): 170–184.
- Goodman, John C., Gerald L. Musgrave, and Devon M. Herrick. 2004. *Lives at risk: Single-payer national health insurance around the world*. Lanham: Rowman & Littlefield.
- Graber, Abraham. 2012. Medusa’s Gaze reflected: A Darwinian dilemma for anti-realist theories of value. *Ethical Theory and Moral Practice* 15: 589–601.
- Graham, Robert. 2004. *Anarchism: A documentary history of libertarian ideas*, vol. 1. Castroville: Black Rose Books.
- Gruber, Jonathan. 2008. Covering the uninsured in the United States. *Journal of Economic Literature* 46(3): 571–606.
- Gryz, Jarek. 2011. On the relationship between the aretaic and the deontic. *Ethical Theory and Moral Practice*. 14: 493–501.
- Guyer, P. 1993. *Kant and the experience of freedom*. Cambridge: Cambridge University Press.
- Guyer, P. 2000. *Kant on freedom, law, and happiness*. Cambridge: Cambridge University Press.
- Habermas, Jürgen. 1990a. *Moral consciousness and communicative action*. Cambridge: The MIT Press.
- Habermas, Jürgen. 1990b. *The philosophical discourse of modernity: Twelve lectures*. Cambridge: The MIT Press.
- Habermas, Jürgen. 1999. Hermeneutic and analytic philosophy. Two complementary versions of the linguistic turn? In *German philosophy since Kant*, ed. Anthony O’Hear. Cambridge: Cambridge University Press.
- Habermas, Jürgen. 2011. Reconciliation through the public use of reason: Remarks on John Rawls’s political liberalism. In *Habermas and Rawls: Disputing the political*, ed. James Gordon Finlayson and Fabian Freyenhagen. New York: Routledge.
- Hacker-Wright, John. 2012. Ethical naturalism and the constitution of agency. *The Journal of Value Inquiry* 46: 13–23.
- Hampton, Jean E. 1998. *The authority of reason*. Cambridge: Cambridge University Press.
- Hare, R.M. 1952. *The language of morals*. Oxford: Oxford University Press.
- Hare, R.M. 1965. *Freedom and reason*. New York: Oxford University Press.
- Hare, R.M. 1981. *Moral thinking: Its levels, method and point*. Oxford: Oxford University Press.
- Hare, R.M. 1996. Off on the wrong foot. In *On the relevance of metaethics: New essays on metaethics*, ed. Jocelyne Couture and Kai Nielsen. Calgary: University of Calgary Press.
- Hare, R.M. 1997. *Sorting out ethics*. Oxford: Oxford University Press.
- Harman, Gilbert H. 1967. Toward a theory of intrinsic value. *Journal of Philosophy* 64(23): 792–804.

- Harrington, Scott E. 2010. The health insurance reform debate. *The Journal of Risk and Insurance* 77(1): 5–38.
- Hauser, Marc D. 2006. *Moral minds: How nature designed our universal sense of right and wrong*. New York: HarperCollins Publishers.
- Hegel, G.W.F. 1991. *Elements of the philosophy of right*. Cambridge: Cambridge University Press.
- Heinaman, Robert. 1993. Rationality, Eudaimonia and Kakodaimonia in Aristotle. *Phronesis* 38 (1): 31–56.
- Hill, T. 1992. *Dignity and practical reason in Kant's moral theory*. Ithaca: Cornell University Press.
- His Holiness the XIV Dalai Lama. *Ethics for our Time*. University of California Television. September 2009.
- Holub, Robert C. 1991. *Jürgen Habermas: Critic in the public sphere*. London: Routledge.
- Horkheimer, Max, and Theodor W. Adorno. 2002. *Dialectic of enlightenment: Philosophical fragments*. Stanford: Stanford University Press.
- Hudson, W.D. 1969. *The is-ought question: A collection of papers on the central problem in moral philosophy*. London: Macmillan.
- Hume, David. 1882. *A treatise on human nature*. London: Longmans, Green, and Co.
- Hursthouse, Rosalind. 1997. Virtue theory and abortion. In *Virtue ethics*, ed. Roger Crisp and Michael Slote, 217–239. Oxford: Oxford University Press.
- Hursthouse, Rosalind. 1999. *On virtue ethics*. Oxford: Oxford University Press.
- Hursthouse, Rosalind. 2011. What does the Aristotelian *Phronimos* know? In *Perfecting virtue: New essays on Kantian ethics and virtue ethics*, ed. Lawrence Jost and Julian Wuerth, 38–57. Cambridge: Cambridge University Press.
- Husserl, Edmund. 1970. *The crisis of European sciences and transcendental phenomenology*. Evanston: Northwestern University Press.
- Jay, Martin. 1973. *The dialectical imagination: A history of the Frankfurt School and the Institute of Social Research, 1923–1950*. Berkeley: University of California Press.
- Kagan, Shelly. 1998. Rethinking intrinsic value. *Journal of Ethics* 2(4): 277–297.
- Kant, Immanuel. 1997. *Groundwork of the metaphysics of morals*. Cambridge: Cambridge University Press.
- Keown, Damien. 2001. *The nature of Buddhist ethics*. New York: Palgrave.
- Kesebir, Seln. 2012. The superorganism account of human sociality: How and when human groups are like beehives. *Personality and Social Psychology Review* 16: 233–261.
- Kleingeld, P. 1995. What do the virtuous hope for? Re-reading Kant's doctrine of the highest good. In *Proceedings of the eighth international Kant congress*, vol. 1, ed. H. Robinson, 91–112. Milwaukee: Marquette University Press.
- Korsgaard, Christine M. 1983. Two distinctions in goodness. *Philosophical Review* 92(2): 169–195.
- Korsgaard, Christine M. 1996. *The sources of normativity*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 2008a. Aristotle's function argument. In *The constitution of agency: Essays on practical reason and moral psychology*, ed. Christine M. Korsgaard. Oxford: Oxford University Press.
- Korsgaard, Christine M. 2008b. From duty and for the sake of the noble: Kant and Aristotle on morally good action. In *The constitution of agency: Essays on practical reason and moral psychology*, ed. Christine M. Korsgaard. Oxford: Oxford University Press.
- Kugihara, N. 2001. Effects of aggressive behaviour and group size on collective escape in an emergency: A test between a social identity model and deindividuation theory. *British Journal of Social Psychology* 40(Pt 4): 575–598.
- Lavoie, Amy. 2008. Hauser presents theory of 'humaniqueness.' *Harvard Gazette*. 14 Feb. 2008. 13 April 2009. <http://news.harvard.edu/gazette/story/2008/02/hauser-presents-theory-of-humaniqueness>

- Lemos, N.M. 2005. The concept of intrinsic value. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 17–31. Dordrecht: Springer.
- Lenz, John R. 2011. How Epicurean science saves humanity in Lucretius. In *Lucretius: His continuing influence and contemporary relevance*, ed. T. Madigan and D.B. Suits, 89–103. Rochester: RIT Press.
- Lillehammer, Hallvard. 2000. The doctrine of internal reasons. *The Journal of Value Inquiry* 34 (4): 507–516.
- Loidolt, Sophie. 2012. A phenomenological ethics of the absolute ought: Investigating Husserl's unpublished ethical writings. In *Ethics and phenomenology*, ed. Mark Sanders and J. Jeremy Wisniewski, 1–37. Lanham: Rowman & Littlefield.
- Macintyre, Alasdair. 1984. *After virtue: A study in moral theory*. Notre Dame: University of Notre Dame Press.
- Macintyre, Alasdair. 1988. *Whose justice, which rationality?* Notre Dame: University of Notre Dame Press.
- Macintyre, Alasdair. 1991. *Three rival versions of moral inquiry: Encyclopedia, genealogy, and tradition*. Notre Dame: University of Notre Dame Press.
- Macintyre, Alasdair. 1999. *Dependent rational animals: Why human beings need the virtues*. Chicago: Carus Publishing Company.
- Mackie, J.L. 1977. *Ethics: Inventing right and wrong*. London: Pelican Books.
- Magnell, Thomas. 1993a. Evaluations as assessments, Part I: Properties and their signifiers. *The Journal of Value Inquiry* 27: 1–11.
- Magnell, Thomas. 1993b. Evaluations as assessments, Part II: Distinguishing assertions and instancing good of a kind. *The Journal of Value Inquiry* 27: 151–163.
- Magnell, Thomas. 2011. The correlativity of rights and duties. *The Journal of Value Inquiry* 45(1): 1–12.
- Magnell, Thomas. 2012. Critical points for civilization, intelligence, and value. *The Journal of Value Inquiry* 46(1): 1–12.
- McCarthy, Thomas. 1978. *The critical theory of Jürgen Habermas*. Cambridge: The MIT Press.
- McDonald, Fritz J. 2010. Agency and responsibility. *The Journal of Value Inquiry* 44(2): 199–207.
- McDowell, John. 1997. Virtue and reason. In *Virtue ethics*, ed. Roger Crisp and Michael Slote, 141–163. Oxford: Oxford University Press.
- McDowell, John. 1998. Values and secondary qualities. In *Mind, value, & reality*, ed. John McDowell, 131–151. Cambridge: Harvard University Press.
- Mill, John Stuart. 2002. *Utilitarianism*. Indianapolis: Hackett Publishing Co.
- Mill, John Stuart. 2006. *On liberty*. Upper Saddle River: Longman.
- Mischel, Walter, and Yuichi Shoda. 1995. A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review* 102(2): 246–268.
- Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nagel, Thomas. 1970. *The possibility of Altruism*. Princeton: Princeton University Press.
- Nagel, Thomas. 1991. *Equality and partiality*. Oxford: Oxford University Press.
- Nagel, Thomas. 1997. *The last word*. Oxford: Oxford University Press.
- Nathanson, Stephen. 2007. Can terrorism be morally justified? In *Social ethics: Morality and social policy*, ed. Thomas A. Mappes and Jane S. Zembaty, 322–332. New York: McGraw Hill.
- Nietzsche, Friedrich. 2006. *Thus spoke Zarathustra*. Cambridge: Cambridge University Press.
- Nietzsche, Friedrich. 2009. *Twilight of the idols: Or how to philosophize with a hammer*, 2009. Oxford: Oxford University Press.
- Nozick, Robert. 1974. *Anarchy, state, and utopia*. New York: Basic Books.
- Nozick, Robert. 1993. *The nature of rationality*. Princeton: Princeton University Press.
- Oderberg, David S. 2007. *Real essentialism*. New York: Routledge.
- O'Hagan, Emer. 2004. Practical identity and the constitution of agency. *The Journal of Value Inquiry* 38(2): 49–59.

- O'Neill, Onora. 1989. *Constructions of reason: Explorations of Kant's practical philosophy*. Cambridge: Cambridge University Press.
- O'Neill, Onora. 1996. *Towards justice and virtue: A constructive account of practical reasoning*. Cambridge: Cambridge University Press.
- Parfit, Derek. 1986. *Reasons and persons*. Oxford: Oxford University Press.
- Parfit, Derek. 2011a. *On what matters*, vol. I. Oxford: Oxford University Press.
- Parfit, Derek. 2011b. *On what matters*, vol. II. Oxford: Oxford University Press.
- Phillips, David. 2001. Gert, Sidgwick, and hybrid theories of rationality. *The Journal of Value Inquiry* 35(4): 439–448.
- Prior, Arthur N. 1949. *Logic and the basis of ethics*. Oxford: Oxford University Press.
- Prior, William J. 2001. Eudaimonism and virtue. *The Journal of Value Inquiry* 35(3): 325–342.
- Punzo, Vincent C. 1980. The normative function of reason as reflectivity: An alternative to Hare's prescriptivism. *The Review of Metaphysics* 33(3): 539–613.
- Rabinowicz, W., and T. Rønnow-Rasmussen. 2003. Tropic of value. *Philosophy and Phenomenological Research* 66: 389–403.
- Rabinowicz, W., and T. Rønnow-Rasmussen. 2005. A distinction in value: Intrinsic and for its own sake. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 17–31. Dordrecht: Springer.
- Railton, Peter Albert. 2003a. Moral realism. In *Facts, values, and norms: Essays toward a morality of consequence*, ed. Peter Albert Railton, 3–42. Cambridge: Cambridge University Press.
- Railton, Peter Albert. 2003b. Facts and values. In *Facts, values, and norms: Essays toward a morality of consequence*, ed. Peter Albert Railton, 43–68. Cambridge: Cambridge University Press.
- Rawls, John. 1971. *A theory of justice*. Cambridge: Harvard University Press.
- Regis, Edward. 1984. *Gewirth's ethical rationalism: Critical essays with a reply by Alan Gewirth*. Chicago: The University of Chicago Press.
- Rehg, William. 1994. *Insight and solidarity: A study in the discourse ethics of Jürgen Habermas*. Berkeley: University of California Press.
- Reitan, Eric. 2006. Self-defense and the principle of generic consistency: Must Gewirth be an absolute pacifist? *Social Theory and Practice* 32(3): 415–438.
- Richards, Robert J. 1993. Birth, death, and resurrection of evolutionary ethics. In *Evolutionary ethics*, ed. Matthew H. Nitecki and Doris V. Nitecki, 113–131. Albany: State University of New York Press.
- Rist, John M. 2002. *Real ethics: Reconsidering the foundations of morality*. Cambridge: Cambridge University Press.
- Roche, Timothy D. 1992. In defense of an alternative view of the foundation of Aristotle's moral theory. *Phronesis* 37(1): 46–84.
- Rorty, Richard. 1979. *Philosophy and the mirror of nature*. Princeton: Princeton University Press.
- Ross, W.D. 2002. *The right and the good*. Oxford: Oxford University Press.
- Rothbard, Murray. 2002. *For a new liberty: The libertarian manifesto*. New York: Collier Macmillan Publishers.
- Rottschaefer, William A. 1998. *The biology and psychology of moral agency*. Cambridge: Cambridge University Press.
- Sartre, Jean-Paul. 2007. *Existentialism is a humanism*. New Haven: Yale University Press.
- Scanlon, T.M. 2000. *What we owe to each other*. Cambridge: Harvard University Press.
- Scanlon, T.M. 2008. *Moral dimensions: Permissibility, meaning, blame*. Cambridge: Harvard University Press.
- Schroeder, Mark. 2008. What is the Frege-Geach problem? *Philosophy Compass* 3(4): 703–720.
- Schurz, Gerhard. 1997. *The is-ought problem: An investigation in philosophical logic*. Dordrecht: Kluwer Academic.
- Seanor, Douglas, and N. Fotion. 1988. *Hare and critics: Essays on moral thinking, with comments by R. M. Hare*. Oxford: Oxford University Press.
- Searle, John R. 1964. How to derive 'ought' from 'is'. *The Philosophical Review* 73(1): 43–58.

- Searle, John R. 1995. *The construction of social reality*. New York: The Free Press.
- Sen, Amartya. 2002. *Rationality and freedom*. Cambridge: Harvard University Press.
- Sen, Amartya. 2009. *The idea of justice*. Cambridge: Harvard University Press.
- Sen, Amartya, and Bernard Williams. 1999. *Utilitarianism and beyond*. Cambridge: Cambridge University Press.
- Sherman, Nancy. 1997. *Making a necessity of virtue: Aristotle and Kant on virtue*. Cambridge: Cambridge University Press.
- Silber, J. 1959. Kant's conception of the highest good as immanent and transcendent. *The Philosophical Review* 68: 469–492.
- Singer, Marcus George. 1961. *Generalization in ethics*. New York: Alfred A. Knopf, Inc.
- Singer, Peter. 1973. The triviality of the debate over 'is-ought' and the definition of 'moral'. *American Philosophical Quarterly* 10(1): 51–56.
- Singer, Marcus George. 2000. Gewirth, Beyleveld, and dialectical necessity. *Ratio Juris* 13(2): 177–195.
- Slobodkin, Lawrence. 1993. The complex questions relating evolution to ethics. In *Evolutionary ethics*, ed. Matthew H. Nitecki and Doris V. Nitecki, 337–347. Albany: State University of New York Press.
- Slote, Michael. 1974. *Metaphysics and essence*. Oxford: Basil Blackwell.
- Slote, Michael. 1992. *From morality to virtue*. Oxford: Oxford University Press.
- Slote, Michael. 1997. Agent-based virtue ethics. In *Virtue ethics*, ed. Roger Crisp and Michael Slote, 219–236. Oxford: Oxford University Press.
- Slote, Michael. 2011. The problem we all have with deontology. In *Perfecting virtue: New essays on Kantian ethics and virtue ethics*, ed. Lawrence Jost and Julian Wuerth, 260–271. Cambridge: Cambridge University Press.
- Smith, Michael. 1994. *The moral problem*. Malden: Blackwell Publishing Ltd.
- Smith, Michael. 2004a. Internal reasons. In *Ethics and the A Priori: Selected essays on moral psychology and meta-ethics*, ed. Michael Smith, 17–42. Cambridge: Cambridge University Press.
- Smith, Michael. 2004b. Internalism's wheel. In *Ethics and the A Priori: Selected essays on moral psychology and meta-ethics*, ed. Michael Smith, 318–342. Cambridge: Cambridge University Press.
- Smith, Michael. 2004c. Objectivity and moral realism: On the significance of the phenomenology of moral experience. In *Ethics and the A Priori: Selected essays on moral psychology and meta-ethics*, ed. Michael Smith, 234–258. Cambridge: Cambridge University Press.
- Snow, Nancy E. 2010. *Virtue as social intelligence: An empirically grounded theory*. New York: Routledge.
- Specter, Matthew G. 2009. Habermas's political thought, 1984–1996: A historical interpretation. *Modern Intellectual History* 6(1): 91–119.
- Specter, Matthew G. 2010. *Habermas: An intellectual biography*. Cambridge: Cambridge University Press.
- Starr, Paul. 2011. *Remedy and reaction: The peculiar American struggle over health care reform*. New Haven: Yale University Press.
- Stohr, Karen. 2006. Contemporary virtue ethics. *Philosophy Compass* 1(1): 22–27.
- Street, Sharon. 2006. A Darwinian dilemma for realist theories of value. *Philosophical Studies* 127: 109–166.
- Sullivan, R. 1989. *Immanuel Kant's moral theory*. Cambridge: Cambridge University Press.
- Swanton, Christine. 2003. *Virtue ethics: A pluralistic view*. Oxford: Oxford University Press.
- Tam, K.-P., et al. 2012. Intersubjective model of value transmission: Parents using perceived norms as reference when socializing children. *Personality and Social Psychology Bulletin* 38: 1041–1052.
- Tännsjö, Torbjörn. 1999. A concrete view of intrinsic value. *The Journal of Value Inquiry* 33(4): 531–535.
- Teichmann, Roger. 2011. *Nature, reason, and the good life*. New York: Oxford University Press.

- Thomson, Judith Jarvis. 1990. *The realm of rights*. Cambridge: Harvard University Press.
- Thomson, Judith Jarvis. 1997. The right and the good. *Journal of Philosophy* 94(6): 273–298.
- Triplett, Timm. 2011a. Justifying morality, Part I: Bernard Gert’s justification. *The Journal of Value Inquiry* 45: 299–308.
- Triplett, Timm. 2011b. Justifying morality, Part II: Beyond justification as clarification. *The Journal of Value Inquiry* 45: 403–417.
- Tuozzo, Thomas M. 1995. Aristotle’s theory of the good and its causal basis. *Phronesis* 40(3): 293–314.
- Uniaque, Suzanne. 1994. *Permissible killing: The self-defence justification of Homicide*. Cambridge: Cambridge University Press.
- Upton, Candace L. 2009. *Situational traits of character: Dispositional foundations and implications for moral psychology and friendship*. Lanham: Lexington Books.
- Veatch, Henry Babcock. 2003. *Rational man: A modern interpretation of Aristotelian ethics*. Indianapolis: Liberty Fund.
- Walton, Douglas. 2009. *Ethical argumentation*. Lanham: Lexington Books.
- Warkentin, Traci. 2009. Whale agency: Affordances and acts of resistance in captive environments. In *Animals and agency: An interdisciplinary exploration*, ed. Sarah E. McFarland and Ryan Hediger, 23–45. Boston: Brill.
- White, Stephen K. 1982. On the normative structure of action: Gewirth and Habermas. *The Review of Politics* 44(2): 282–301.
- Williams, Bernard. 2006. *Ethics and the limits of philosophy*. Abingdon: Routledge.
- Wittgenstein, Ludwig. 1965. A lecture on ethics. *The Philosophical Review* 74: 3–12.
- Zagzebski, Linda Trinkaus. 1996. *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press.
- Zimmerman, M.J. 2005a. Defending the concept of intrinsic value. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 153–168. Dordrecht: Springer.
- Zimmerman, M.J. 2005b. Intrinsic value and individual worth. In *Recent work on intrinsic value*, ed. Toni Rønnow-Rasmussen and Michael J. Zimmerman, 191–205. Dordrecht: Springer.