

Theory and Decision Library A: Rational Choice in Practical
Philosophy and Philosophy of Science

John R. Welch

Moral Strata

Another Approach to Reflective
Equilibrium

 Springer

Theory and Decision Library A:

Rational Choice in Practical Philosophy
and Philosophy of Science

Volume 49

Series editor

Julian Nida-Rümelin

Universität München, Munich, Berlin, Germany

This series deals with practical and social philosophy and also foundational issues in philosophy of science in general that rely on methods broadly based on rational choice. The emphasis in the Series A is on well-argued, thoroughly analytical and philosophical rather than advanced mathematical treatments that use methods from decision theory, game theory and social choice theory. Particular attention is paid to work in practical philosophy broadly conceived, the theory of rationality, issues in collective intentionality, and philosophy of science, especially interdisciplinary approaches to social sciences and economics.

Assistant Editor: Martin Rechenauer (München)

Editorial Board: Raymond Boudon (Paris), Mario Bunge, (Montréal), Franz Dietrich, (Paris & East Anglia), Stephan Hartmann, (Tilburg), Martin van Hees (Amsterdam), Isaac Levi (New York), Richard V. Mattessich (Vancouver), Bertrand Munier (Cachan), Olivier Roy (Bayreuth), Amartya K. Sen (Cambridge), Brian Skyrms, (Irvine), Wolfgang Spohn (Konstanz), Katie Steele, (London School of Economics).

More information about this series at <http://www.springer.com/series/6616>

John R. Welch

Moral Strata

Another Approach to Reflective
Equilibrium

John R. Welch
Saint Louis University – Madrid Campus
Madrid
Spain

ISSN 0921-3384 ISSN 2352-2119 (electronic)
ISBN 978-3-319-08012-3 ISBN 978-3-319-08013-0 (eBook)
DOI 10.1007/978-3-319-08013-0
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014945757

© 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)

These heavy sands are language
—James Joyce, *Ulysses*

Preface

This book began an embarrassingly long time ago. In looking for discussion material for an ethics class, I chanced across my copy of William Styron's *Sophie's Choice*. The work's central dilemma intrigued my students, who concluded that an act-utilitarian and a Kantian using the second formulation of the categorical imperative would make very different decisions about the case. The question was then unavoidable: Which theory offers the better ethical advice? I was dissatisfied with my own attempt to answer.

Meanwhile, I had become accustomed to responding to students' contrasts of science and ethics by citing Alan Gewirth's "Positive 'Ethics' and Normative 'Science,'" which warns of the fallacy of disparateness: "the fallacy of discussing one field on one level or in one respect and the other field on a quite different level or in a quite different respect." To contrast a scientific discovery such as the molecular structure of DNA with an ethical discussion of the highest good is an instance of this fallacy. Yet I was unable to be clearer about Gewirth's levels or respects until I discovered Larry Laudan's analysis of scientific discourse in *Science and Values*. In reading the work, I formed the hypothesis that moral discourse, like scientific discourse, could be analyzed into factual, methodological, and axiological levels. This hypothesis gained momentum when, in the book's Epilogue, Laudan himself mentioned the possibility of extending his approach to moral theory.

Rationality in morality, I thought, appears to be governed by the cognitive ideal of reflective equilibrium among levels of moral discourse analogous to Laudan's levels of scientific discourse. I proposed this ideal in "Science and Ethics: Toward a Theory of Ethical Value," which can be seen as a kind of mission statement for this book. But the article offered only the sketchiest indications of how reflective equilibrium might be attained at each of these levels, and very difficult technical problems lay half-submerged in each case. My attempts to resolve these problems led to explorations of quantitative inductive logics and comparative decision theory.

In the course of these explorations, I began to see what I take to be rational grounds for choice among theories. Rival ethical theories can offer conflicting advice about dilemmas, and quantitative inductive logics can be used to resolve a common sort of dilemma. Whenever this occurs, any theory that recommends the inductively preferred option secures an advantage over theories that recommend

other options. In addition, decision theory can be employed to guide the choice between one theory and another, particularly when formulated in terms of comparative plausibilities and utilities. This book presents such a version of decision theory, offered in the hope that it will aid in the quest for reflective equilibrium.

Quotation marks in this work are handled as follows. Double quotes are employed for short quotations, whether attributed or not, and quotations within long quotations. No quotes are used for long quotations, which are set off from their context by indents and smaller type. Single quotes are used for quotations within short quotations and words cited as words, such as the predicate 'just'.

The development of the outlook presented in this volume was facilitated by interactions with many people: students, colleagues, conference participants, anonymous reviewers, and editors, among others. These interlocutors are too numerous to be listed individually, but I cannot fail to mention my parents, the late Mary V. and Robert J. Welch, who showed by example the centrality of morality to human life. Nor could I omit my colleagues Renzo Llorente, Olga Ramírez Calle, and Jawara Sanford, who commented insightfully on sundry parts of the manuscript. Talented people at Springer who played vital roles in this project include Associate Editor Lucy Fleet, Assistant Editor Martin Rechenauer, and Senior Editorial Assistants Diana Nijenhuijzen and Mireille van Kan. Finally, my wife Cristina and son Guillermo formed the uniquely supportive environment that enabled this work to be completed. Each, in different ways, has helped me through this project. It is an unmingled pleasure to thank you all.

Contents

1	Discursive Strata	1
1.1	Moral Strata	1
1.2	Origins of Reflective Equilibrium	3
1.3	Problems with Reflective Equilibrium	5
1.3.1	Moral Conservatism	5
1.3.2	Moral Diversity	6
1.3.3	The Moral Weight of Considered Judgments	8
1.3.4	The Nature of Considered Judgments	9
1.3.5	Intuitionism	10
1.4	A Proposal for Wide Reflective Equilibrium	11
1.5	Conclusion	14
	References	15
2	Saving the Moral Phenomena	17
2.1	Inductive Molding	17
2.2	Core Classification	19
2.3	Core Classification in Ethics	21
2.3.1	Prototype Theory	21
2.3.2	Washington's Cherry Tree	22
2.4	A Standard for Analogy	25
2.4.1	Inductive Cogency	26
2.4.2	Analogy as Induction	28
2.4.3	Analogy and Inductive Strength	30
2.5	Applying the Standard to Ethical Analogies	36
2.5.1	Clash Points	36
2.5.2	The Grain Merchant	37
2.6	Conclusion	41
	References	43
3	Comparative Decision Theory	47
3.1	Toward a Realistic Decision Theory	47
3.2	How to Choose a Theory	49

3.3	Decisions under Risk	54
3.3.1	The Basics	54
3.3.2	Adapting the Basics	57
3.4	Relative Disutility	64
3.4.1	The Hintikka-Pietarinen Proposal	64
3.4.2	Generalizing the Hintikka-Pietarinen Proposal	65
3.5	Comparative Decision Theory	68
3.5.1	The Basic Binary Case	68
3.5.2	The Full Binary Case	72
3.5.3	The Finite General Case	76
3.6	Shoring Up the Foundations	78
3.6.1	Transitivity	78
3.6.2	The Principle of Independence	81
3.6.3	Suspending Judgment	87
3.7	Conclusion	89
	References	90
4	Working with Moral Means	95
4.1	Moral Instrumentality	95
4.2	Distinguishing Means	96
4.3	Means, Ends, and Their Critics	97
4.4	Instrumental Moral Sentences	99
4.5	Practical Inference	101
4.5.1	The Practical Syllogism	101
4.5.2	Assessing Practical Inference	106
4.5.3	Chaining Practical Inferences	113
4.5.4	The Kantian Alternative	113
4.6	Moral Theory Choice	117
4.6.1	Plausibility	118
4.6.2	Utility	123
4.6.3	Plausibilistic Expectation	126
4.7	Conclusion	129
	References	129
5	Securing Our Moral Ends	133
5.1	Moral Teleology	133
5.2	Distinguishing Ends	134
5.3	Justifying Ends	135
5.4	Teleological Moral Descriptions	138
5.5	Teleological Moral Directives	141
5.5.1	Plausibility	143
5.5.2	Utility	144
5.5.3	Plausibilistic Expectation	145
5.6	Amplifying Mixed Deontology	147
5.7	Morality as an End	148

5.8	Why Coherence?	155
5.8.1	Probabilistic Coherence	155
5.8.2	Plausibilistic Coherence	158
5.8.3	Coherence in Contemporary Epistemology	165
5.9	Conclusion	166
	References	166
6	Remedies for Reflective Disequilibrium	169
6.1	Reflective Disequilibrium	169
6.2	Phenomenal Disequilibrium	171
6.3	Instrumental Disequilibrium	176
6.4	Teleological Disequilibrium	178
6.5	Extra-Moral Disequilibrium	180
6.5.1	Supererogation	181
6.5.2	Moral Obligation	186
6.5.3	Overridingness	190
6.6	Conclusion	194
	References	194
Index		197

Chapter 1

Discursive Strata

Abstract Chapter 1 introduces the strata that structure this work: phenomenal, instrumental, and teleological moral discourse. After an overview of the approach to moral strata presented in this volume, the chapter offers a quick reprise of the method of reflective equilibrium as elaborated by Goodman, Rawls, and Daniels. It then considers five objections to this method in its canonical formulations. Objections couched in terms of moral conservatism, moral diversity, and the moral weight of considered judgments are judged unsuccessful, while objections based on the nature of considered judgments and the relation to intuitionism are found to be more problematic. In order to meet these last two objections, the chapter reworks the received view of reflective equilibrium by defining an alternative notion of wide reflective equilibrium. This alternative is presented as a cognitive ideal: coherence among phenomenal, instrumental, and teleological discursive strata in addition to background theories. How moral discourse might achieve coherence of this sort is the subject of successive chapters.

1.1 Moral Strata

Where goals are concerned, the clearer the better. This book is an attempt to clarify the goal of moral inquiry. It undertakes this project in two stages. Its opening chapter proposes that the goal of moral inquiry is a specific kind of reflective equilibrium. Subsequent chapters then suggest ways of reaching (or at least approximating) this goal.

To specify the requisite form of reflective equilibrium, the book invokes the notion of discursive strata. In any field, discursive strata are formed by sentences and differentiated by function. Sentences that perform the same kind of linguistic function clump together, so to speak, to form a stratum. Where the field is moral inquiry, three linguistic functions appear to be central. In the morally phenomenal stratum, predicates like ‘dishonest’, ‘loyal’, ‘cowardly’, ‘generous’, and ‘cruel’ are applied to actions, policies, and persons; ‘That took courage’ is an instance. The morally instrumental stratum concentrates on means for attaining moral ends, as in ‘Justice requires an independent judiciary’. The correlative stratum is morally teleological, formed by sentences that focus on moral ends like ‘Seek the greatest happiness of the greatest number’.

The book's opening chapter recruits these strata to delineate the appropriate form of reflective equilibrium. The result is a version of wide reflective equilibrium here proposed as a goal for moral inquiry. How this goal might be achieved, stratum by stratum, is the subject of the rest of the work.

Chapter 2 addresses the phenomenal stratum. Its initial step is to introduce a fundamental form of classification called 'core classification'. To develop this concept, the chapter formulates the analogy thesis: core classification in general, and moral core classification with terms like 'honest' in particular, is carried out by analogy. In addition, the chapter holds that good analogies can be distinguished from bad ones by appeal to a standard of inductive cogency. This standard is specified with the help of quantitative inductive logics in the tradition of Carnap, Hintikka, Kuipers, and Niiniluoto. The chapter shows how the use of these logics can reduce vagueness and provide an in-principle solution to morally phenomenal disagreements. It illustrates these claims through an extended discussion of the moral dilemma faced by Cicero's grain merchant.

Chapter 3 postpones the treatment of instrumental and teleological moral strata in order to marshal technical resources needed in subsequent chapters. The chapter opens with a survey of four approaches to theory choice. The result of this survey is an endorsement of decision theory as a guide to theory choice. However, decision theory has a serious problem of numeric poverty: standard applications of decision theory require point-valued utilities (or point-valued probabilities and utilities), but we rarely have precise and reliable values for these inputs. Consequently, this chapter pleads for a more widely applicable form of decision theory. It proceeds to argue that a comparative version of decision theory fills the bill. It shows that many choices among theories can be based on merely comparative plausibilities and utilities.

Chapter 4 analyzes the instrumental stratum of moral discourse. Since this stratum is composed of substrata that can be demarcated in different ways, the chapter considers three different groupings: individual sentences, inferences, and theories. It contends that individual sentences can be confirmed or disconfirmed through scrupulous observation and inductive logic. The chapter also treats the justifiability of practical inferences, proposing that they be evaluated by the standard of inductive cogency introduced in Chap. 2. Finally, the chapter grapples with the issue of moral theory choice. It holds that a moral theory can be chosen on instrumental grounds by applying the comparative decision theory of Chap. 3. To illustrate the procedure, it undertakes a comparative evaluation of Kantian, Benthamite, and Frankenan theories as applied to Sophie's choice.

Chapter 5 investigates the teleological stratum of moral discourse. It distinguishes teleological descriptions such as 'The highest good is the greatest happiness of the greatest number' from teleological directives like 'Act from the good will'. The chapter maintains that teleological descriptions can be viewed as hypotheses. As such, they can be confirmed or disconfirmed through hypothetico-deductive reasoning analogous to that employed in the sciences. In addition, the chapter shows that choice among teleological directives can be reasonably guided by the comparative decision theory of Chap. 3. Finally, the chapter ponders the possibility of justifying the higher-order end of being moral. It urges that coherence requires accepting

‘I ought to be moral’, where coherence is understood in terms of wide reflective equilibrium.

The concluding Chap. 6 proposes remedies for a common affliction: reflective disequilibrium. This affliction can result from inconsistencies within moral strata or between moral and nonmoral discourse. The chapter claims that reflective disequilibrium within the phenomenal stratum can be relieved in some cases through recourse to the standard of inductive cogency defined in Chap. 2. Reflective disequilibrium within the instrumental stratum may be resolved intra-theoretically, as illustrated by the application of Frankonian theory to the case of *United States v. Holmes*, or finessed inter-theoretically through ascent to the teleological stratum of moral discourse. Reflective disequilibrium within the teleological stratum can often be reduced inter-theoretically with the help of comparative decision theory or intra-theoretically on consequentialist grounds. Finally, disequilibrium between moral and nonmoral discourse can be reduced by hewing to a modest version of the overridingness thesis.

1.2 Origins of Reflective Equilibrium

C. S. Peirce famously distinguished two forms of thought: “Thought in action has for its only possible motive the attainment of thought at rest” ([1878] 1986, p. 263). Thought in action is the activity of inquiry; thought at rest is inquiry’s goal. Thought at rest, according to Peirce, is settled belief. So inquiry aims at settled belief.

The view that the goal of inquiry is settled belief has been held in one form or another by a number of recent philosophers, notably those influenced by Nelson Goodman.¹ Its best-known restatement is John Rawls’ concept of reflective equilibrium, routinely described as the most widely used method in contemporary ethics (e.g., Varner 2012, p. 11). Attaining reflective equilibrium, according to Rawls, requires the “mutual adjustment of principles and considered judgments” (1971, p. 20 n. 7). Even though this description is rooted in Rawls’ discussion of justice, he remarks at least twice that the process of mutual adjustment of principles and considered judgments is “not peculiar to moral philosophy” (1971, p. 20 n. 7, 49), and he cites Goodman’s observations about the justification of general rules and particular inferences in both deductive and inductive logic: “The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either” (Goodman 1979, p. 64).

The process of justification Goodman refers to is not peculiar to logic. Russell and Whitehead, for example, make cognate remarks about justification in mathematics:

¹ For example, Rawls (1971, p. 20, 46–51), Putnam (1983, pp. 201–202), and Elgin (1989, p. 91, 94; 1996).

[T]he chief reason in favour of any theory on the principles of mathematics must always be inductive, *i.e.* it must lie in the fact that the theory in question [which enunciates principles] enables us to deduce ordinary mathematics [considered mathematical judgments]. In mathematics, the greatest degree of self-evidence is usually not to be found quite at the beginning, but at some later point; hence the early deductions, until they reach this point, give reasons rather for believing the premisses because true consequences follow from them, than for believing the consequences because they follow from the premisses. (Russell and Whitehead 1927, vol. I, p. v)

Similarly, Vann McGee draws the methodological parallel between justification in logic and justification in the empirical sciences:

The methodological moral to be drawn from this [putative counter-examples to *modus ponens*] is that, when we formulate general laws of logic, we ought to exercise the same sort of caution we exercise when we make inductive generalizations [empirical principles] in the empirical sciences. We must take care that the instances [considered empirical judgments] we look at in evaluating a proposed generalization are diverse as well as numerous. (1985, p. 468)

Hence this process of mutual adjustment of principles and considered judgments goes on in moral philosophy, logic, the foundations of mathematics, and the empirical sciences. It is not limited to these contexts, however. It appears to occur wherever rational inquirers employ principles and considered judgments to inquire.

The goal of reflective equilibrium can take more than one form. Rawls' distinction between narrow and wide reflective equilibrium is present in all but name in *A Theory of Justice* (1971, p. 48, 49), and Rawls began to counterpose the two forms explicitly not long after the publication of his best-known work (e.g., 1975, p. 8, 21). Suppose that we undertake a process of mutual adjustment of our principles and considered judgments; if the end result is coherent, we have reached a state of narrow equilibrium. The critical edge of this process is likely to be dull, however. Though we may reject a considered judgment here and revise a principle there, the process is meant to salvage as much of our belief structure as possible. Suppose, on the other hand, that we attempt to subject the entire structure of relevant beliefs to philosophical scrutiny, contrasting our native point of view with alternative conceptions. In such cases,

we are interested in what conceptions people would affirm when they have achieved wide and not just narrow reflective equilibrium, an equilibrium that satisfies certain conditions of rationality. That is, adopting the role of observing moral theorists, we investigate what principles people would acknowledge and accept the consequences of when they have had an opportunity to consider other plausible conceptions and to assess their supporting grounds. Taking this process to the limit, one seeks the conception, or plurality of conceptions, that would survive the rational consideration of all feasible conceptions and all reasonable arguments for them. We cannot, of course, actually do this, but we can do what seems like the next best thing, namely, to characterize the structures of the predominant conceptions familiar to us from the philosophical tradition, and to work out the further refinements of these that strike us as most promising. (Rawls 1975, p. 8)

Hence wide reflective equilibrium, as Rawls conceives it, emerges only if we escape the narrow circle of our own favored principles and considered judgments and confront them with the widest possible range of alternative conceptions. As Norman

Daniels observes, “We must show why it is reasonable to hold these principles and beliefs, not just that we happen to do so” (1996, p. 1).

Daniels stresses the importance of wide reflective equilibrium for moral theory, but he rejects a “two-tiered view of moral theories” that relies on a set of considered judgments and a set of principles (1979, p. 256). Instead, he proposed a more complex view of moral theories that registers the importance of background theories. For example, Rawls’ conclusions about justice as fairness are derived with the help of a set of background theories that includes “a theory of the person, a theory of procedural justice, general social theory, and a theory of the role of morality in society (including the ideal of a well-ordered society)” (Daniels 1979, p. 260). Hence we want coherence not only among our considered judgments and principles; we also want both to cohere with our background theories. Consequently, Daniels expanded the definition of wide reflective equilibrium as follows: “The method of wide reflective equilibrium is an attempt to produce coherence in an ordered triple of sets of beliefs held by a particular person, namely, (a) a set of considered moral judgments, (b) a set of moral principles, and (c) a set of relevant background theories” (1979, p. 258).

1.3 Problems with Reflective Equilibrium

Like other cognitive ideals, the ideal of reflective equilibrium has been challenged. In this section I want to air five of these challenges. The discussion aims to strike a sort of Aristotelian mean between too much and too little. To say too much would be to shoehorn the rest of the book into this one preliminary section, for one way to look at the book as a whole is as a defense of wide reflective equilibrium. But to say too little would leave the reader with the sense that substantive objections to this cognitive ideal have simply gone unheeded.

1.3.1 *Moral Conservatism*

One complaint about the method of reflective equilibrium is that it is uncritically conservative—a volley returned in the opposite direction by the method’s proponents.² An early advocate of this point of view is Richard Brandt, who worries that the coherence proper to the method “may be no more than a reshuffling of moral

² Rawls targets utilitarianism in particular: “the [utilitarian] choice [of “ideals of the person”] does depend upon existing desires and present social circumstances and their natural continuations into the future.” By contrast, justice as fairness and perfectionism “establish independently an ideal conception of the person and of the basic structure so that not only are some desires and inclinations necessarily discouraged but the effect of the initial circumstances will eventually disappear” (1971, p. 262). Daniels also claims that “utilitarianism is biased toward the status quo” (1996, p. 94).

prejudices” (1979, p. 22). More recently, Gilbert Harman observes “The method is conservative in that we start with our present views and try to make the least change that will best promote the coherence of our whole view” (2003, p. 416). Indeed, Allen Wood rejects the method because some of our beliefs are corrupted by “the radical evil of our social condition,” which would contaminate any equilibrium that retains such beliefs (2008, p. 5).

These concerns appear to be valid for the narrow variety of reflective equilibrium but not for the wide. Wide reflective equilibrium requires the widest possible reflection, which includes all the critical resources we can bring to bear. Ron Amit has compared Rawlsian method to the Frankfurt School’s method of immanent criticism, arguing that it holds out “an idealized and improved mirror against which we confront our own institutions” (2006, p. 182). In addition, Katarzyna de Lazari-Radek and Peter Singer point out that commitment to wide reflective equilibrium includes commitment to scientific beliefs, which means that evolutionary theory could be used “to reject many widely shared moral intuitions” (Lazari-Radek and Singer 2012, p. 30). Consequently, the method does not appear to be inherently conservative.

1.3.2 *Moral Diversity*

Some of these same critics charge that moral diversity derails the quest for reflective equilibrium. Singer puts the point as follows:

If I am right in attributing this version of the reflective equilibrium idea to Rawls, then Rawls is a subjectivist about morality in the most important sense of this often-misused term. That is, it follows from his views that the validity of a moral theory will vary according to whose considered judgments the theory is tested against. There is no sense in which we can speak of a theory being objectively valid, no matter what considered moral judgments people happen to hold. If I live in one society, and accept one set of considered moral judgments, while you live in another society and hold a quite different set, very different moral theories may be “valid” for each of us. There will then be no sense in which one of us is wrong and the other right. (1974, p. 494)

Brandt objects along similar lines:

Moreover, moral intuitions differ from one individual or culture to another. Where one person thinks promise-keeping or sexual taboos are highly important—these beliefs have high initial credence level—and another does not, the search for reflective equilibrium will only produce different moral systems, and offers no way to relieve the conflict. Nor is this matter trivial. Moral disagreement does not exist only between our own reflective equilibria and those of some primitive tribes, or on relatively superficial matters. It exists among sophisticated civilized persons and in core areas.... (1979, p. 22)

Singer and Brandt are actually making four interrelated points: there are different considered judgments; because there are different considered judgments, there are different moral systems; different moral systems can be in reflective equilibrium; and reflective equilibrium will not relieve conflicts among moral systems. I will comment briefly on each in turn.

The claim about differing considered judgments is an empirical point, of course. Brandt (1954) contributed to research on this subject, and I think it likely that, in some carefully qualified sense, the claim is true. Some differences among considered judgments may be attributable to error by one or more parties; others may not. Errors that affect considered judgments would not create methodological problems, for reflective equilibrium requires correcting considered judgments when necessary. But if differing considered judgments cannot be explained away as errors, the proponent of reflective equilibrium will simply accept the differences. Accordingly, the analysis of morally phenomenal discourse in Sect. 2.3.2 explicitly recognizes the possibility of moral prototypes that are culturally variant. I will not assume it impossible for one culture to stress moral properties that are minimized or ignored by another.

However, if two cultures do in fact have different considered judgments, different moral systems would result, just as Singer and Brandt say. Indeed, different considered judgments that did *not* lead to different moral systems would require strenuous explanation. Note, however, that ‘different’ here means ‘non-identical’, not ‘non-overlapping’. Careful analysis may show that different sets of equilibrium beliefs nevertheless overlap. In fact, this is what the historical record would lead us to expect. It shows substantial areas of moral agreement and narrowly delimited areas of moral disagreement, typically over exceptions to agreed-upon moral principles (Rachels 1986, pp. 19–22).

Different sets of moral beliefs, Singer and Brandt imply, can be in reflective equilibrium.³ I suggest, once again, that they are right; “the prospects of divergence in wide reflective equilibrium remain significant” (Daniels 1996, p. 8). But this does not show that there is something wrong with reflective equilibrium. On the contrary, it shows that there is something right with it. Reflective equilibrium is a criterion of justification, not a criterion of truth (Daniels 1979, p. 277; Knight 2006, p. 220). As such, it is fully consistent with the fact that beliefs can be justified yet nonetheless false. Just as we can say that different historical beliefs about the shape of the earth, say, can be justified without committing ourselves to more than one geological truth, we can recognize that different moral beliefs can be justified without committing ourselves to more than one moral truth. That different sets of moral beliefs can be in reflective equilibrium is just what we would expect from a reliable criterion of justification.

Finally, we have Singer’s and Brandt’s point that reflective equilibrium would provide no way to relieve conflicts among different moral systems. Take people in a culture with a collective belief that human sacrifice is necessary to prevent the world from being destroyed, for instance. Thinking people in such a culture might attain reflective equilibrium provided they lack the empirical resources that would show the underlying cosmological belief to be false. Imagine such people being confronted by contemporaries from a second culture who object to human sacrifice on the grounds that it is immoral and causally independent of the world’s survival.

³ Rawls himself mentions this possibility but dismisses it for being “far beyond our reach” (1971, p. 50).

Would there be a way to relieve the conflict? In principle, yes: halt the sacrifices and watch what happens. The watching—short or long—should settle the conflict. Here, then, is a counter-example to the claim that there is no way to relieve conflicts among different moral systems. Obviously, though, this in-principle solution might not work in practice; people in the first culture might refuse to pursue this solution out of fear. I suggest, then, that blanket statements that there is or is not a way to relieve such conflicts are just too coarse. I hope to show, however, that when conflicts are identified as morally phenomenal, instrumental, and teleological, resolution can be sought in rational ways.

The upshot, I think, is that even though much of what Singer and Brandt say is true, none of it spoils the party for reflective equilibrium. Suitably understood, reflective equilibrium is not committed to unanimity of considered judgments, uniqueness of reflective equilibria, or universal resolution of conflicts among moral systems. But the pursuit of reflective equilibrium can, I claim, reduce or eliminate moral conflict in a rational way. The defense of this claim is the work of later chapters.

1.3.3 *The Moral Weight of Considered Judgments*

The moral weight that reflective equilibrium grants to considered judgments has been criticized. Singer, for example, suggests that it would be better to ignore them and work with moral axioms instead:

Why should we not rather make the opposite assumption, that all the particular moral judgments we intuitively make are likely to derive from discarded religious systems, from warped views of sex and bodily functions, or from customs necessary for the survival of the group in social and economic circumstances that now lie in the distant past? In which case, it would be best to forget all about our particular moral judgments, and start again from as near as we can get to self-evident moral axioms. (1974, p. 516)

Similarly, Brandt remarks that considered judgments may have credence but not credibility:

There is a problem here quite similar to that which faces the traditional coherence theory of justification of belief: that the theory claims that a more coherent system of beliefs is better justified than a less coherent one, but there is no reason to think this claim is true unless some of the beliefs are initially credible—and not merely initially believed—for some reason other than their coherence, say, because they state facts of observation. In the case of normative beliefs, no reason has been offered why we should think that initial credence levels, for a person, correspond to credibilities. The fact that a person has a firm normative conviction [considered judgment] gives that belief a status no better than fiction. Is one coherent set of fictions supposed to be better than another? (1979, p. 20)

Two points, I think, can be offered in response. The first is that Brandt (though not Singer) eventually softened his position on considered judgments. In fact, the science-based theory of ethics Brandt came to favor can heed considered judgments. This theory, he says,

will respect “considered opinions” about what is right, at least as warning signs where there is disparity, for in the form these take in ordinary life they seem mostly to occur either as expressions of compassion (being shocked by harm to others) or repugnance at some forms of behavior which members of the society normally have come to find repugnant because of their usual connection with tendencies to harm. The “considered judgments” doubtless have been strongly influenced by the norms of society, but since there seems to be a process of pruning which modifies these norms so as to drop the useless or harmful norms and also some process of thoughtfully addressing new problems, one’s society-produced judgments deserve some respect. Norms have been much influenced by what thoughtful people have wanted in the moral code of the society. (Not always: see the... morality of homosexual behavior.) So it is reasonable to accept a principle of conservatism here. (1990, p. 277)

Secondly, I hope to show in successive chapters that if reflective equilibrium is reconceptualized with respect to phenomenal, instrumental, and teleological strata, discourse within each stratum can not only express belief but acquire credibility. I argue for this conclusion within the phenomenal stratum in Chap. 2; within the instrumental stratum in Chap. 4; and within the teleological stratum in Chap. 5. These arguments are fully consistent with the conviction that justification of moral discourse within each of these strata is an intricate and arduous affair and that the mutual adjustment of these strata is more demanding still. Moral discourse is fallible through and through. But in no stratum, I maintain, can a no-credibility argument be sustained.

1.3.4 *The Nature of Considered Judgments*

The other difficulty with considered judgments is just what they are supposed to be. Rawls defined them as “those judgments in which our moral capacities are most likely to be displayed without distortion” (1971, p. 47). However, Goodman’s seminal contrast was between general rules and particular inferences:

Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences. . . . The point is that rules and particular inferences alike are justified by being brought into agreement with each other. . . . Thus the interplay we observed between rules of induction and particular inductive inferences is simply an instance of this characteristic dual adjustment between definition and usage, whereby the usage informs the definition which in turn guides extension of the usage. (1979, p. 64, 66)

In the view of many readers, Rawls likewise meant to distinguish considered judgments, which are particular, from principles, which are general (e.g., Daniels 1996, p. 1). But Rawls uses the injustice of religious intolerance and racial discrimination as examples of considered judgments in *A Theory of Justice* (1971, p. 19), and the post-*Theory* Rawls says explicitly that considered judgments can have any level of generality: “People have considered judgments at all levels of generality, from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions” (1975, p. 8; cf. Richardson 1994, p. 178). However, this complicates the contrast between considered judgments and principles considerably. Just how are they supposed to differ? We return briefly to this problem in Sect. 1.5.

1.3.5 *Intuitionism*

A final difficulty with the method of reflective equilibrium is its relation to intuitionism.⁴ Some commentators have accused Rawls of espousing a form of subjective intuitionism. R. M. Hare claims “Intuitionism is nearly always a form of disguised subjectivism. Rawls does not call himself an intuitionist; but he certainly is one in the usual sense” (1973, p. 146). Similarly, Brandt includes Rawls among those who espouse “the method of intuitions” (1979, pp. 19–22). In a related line of argument, Stephen Stich draws on the heuristics and biases literature to urge that people under the sway of a fallacious heuristic—an intuition—might adopt some “daffy inferential rule” and yet be in a state of reflective equilibrium (1990, p. 86).

The tendency to link reflective equilibrium with intuitionism is reinforced by a penchant for describing considered judgments as intuitions. Jared Bates, for example, understands the method of reflective equilibrium in epistemology to require testing theories against “our intuitions about cases of justified and unjustified belief” (2004, p. 45). Analogously, David Gauthier takes the method of reflective equilibrium in morality to entail appeal to moral intuitions (that is, considered moral judgments):

If the reader is tempted to object... on the ground that his *moral intuitions* are violated, then he should ask what weight such an objection can have, if morality is to fit within the domain of rational choice. We emphasize the radical difference between our approach, in which we ask what view of social relationships would rationally be accepted *ex ante* by individuals concerned to maximize their utilities, from that of moral coherentists and defenders of “reflective equilibrium,” who allow initial weight to our *considered moral judgements*. (1986, p. 269; emphasis added)

To clarify the relation between reflective equilibrium and intuitionism, let us begin with Rawls’ own definition of intuitionism: “Intuitionist theories, then, have two features: first, they consist of a plurality of first principles which may conflict to give contrary directives in particular types of cases; and second, they include no explicit method, no priority rules, for weighing these principles against one another: we are simply to strike a balance by intuition, by what seems to us most nearly right” (1971, p. 34). Evaluated in the light of this definition, Rawls’ principles of justice do not form an intuitionist system. Even though he admits a plurality of first principles, his first principle of justice has clear priority over the second.

In addition, the method of reflective equilibrium is not intuitionist in a straightforward sense. Intuitionism in the usual sense is a form of foundationalism, and some commentators accordingly take reflective equilibrium to be foundationalist (e.g., Bates 2004, pp. 48–50). However, as Daniels points out, reflective equilibrium is not foundationalist (1979, pp. 264–265, 1996, p. 4). In the quest for reflective equilibrium, none of our beliefs is held immune from revision. Because there

⁴ There are at least four different species of intuitionism: reliabilism, experientialism, reflectionism, and contextualism (Sinnott-Armstrong 2006, p. 186). The reliability of moral intuitions has been attacked (Singer 2005) and defended (Tersman 2008) in the light of recent neuroscience.

is no priority among considered judgments, principles, and background theories, a considered judgment might be rejected because it conflicts with principles or background theories, or a principle might be abandoned because it is inconsistent with considered judgments or background theories. Consequently, since reflective equilibrium is not foundationalist, it is not intuitionism in the usual sense. Reflective equilibrium is arguably not foundationalist in another, more radical sense. Michael DePaul claims that reflective equilibrium and foundationalism “are not really positions on the same topic” because reflective equilibrium is a method, while foundationalism is an epistemic account of beliefs (1986, p. 68).

Nevertheless, some might grant that reflective equilibrium is not intuitionist in the two previous senses but counter that it is in a third. That is, even though Rawls’ principles of justice are prioritized and reflective equilibrium permits neither considered judgments nor principles to be foundational, the method does rely on moral intuitions in cases of reflective disequilibrium. Given the real possibility that common morality is inconsistent (Brand-Ballard 2003, pp. 231–232), reflective disequilibrium is not a peripheral matter. Say, then, that we notice an inconsistency between considered judgments and principles but resolve the inconsistency through an intuition that the principles are more reliable. In another instance, of course, we might appeal to an intuition in favor of the considered judgments. But no matter which way we lean, the argument goes, we are relying on moral intuitions, and these intuitions serve as a cognitive foundation.

One response to this argument is to suggest that those who read Rawls as an intuitionist fail to distinguish between narrow and wide reflective equilibrium (Daniels 1979, p. 267 n. 17). For if the goal is not just coherence of our native belief system but coherence acquired through the widest possible reflection, through radical contrasts of our beliefs with alternative conceptions, the choice of considered judgments over principles (or vice versa) is not due to an intuition; it is due to disciplined philosophical reflection. This response is on the right track, I think, but there is more to be said. If wide reflective equilibrium is suitably reconceptualized, any residual tendency to conflate it with intuitionism vanishes. A proposal along these lines occupies the following section.

1.4 A Proposal for Wide Reflective Equilibrium

Consistent with the view that the quest for reflective equilibrium is not restricted to moral philosophy, a reconceptualization of reflective equilibrium can be extrapolated from the philosophy of science.⁵ Larry Laudan has proposed what he calls the reticulated model of scientific rationality (1984, pp. 50–66). Perhaps the simplest way to introduce its central ideas is by contrast with the older, hierarchical model of justification associated with Popper, Hempel, and Reichenbach, among others. The

⁵ The following proposal builds on Welch (1994, pp. 281–282).