

REFLECTIVE EQUILIBRIUM

Essays in Honour of Robert Heeger

Edited by

WIBREN VAN DER BURG

*Schoordijk Institute,
Tilburg University, The Netherlands*

and

THEO VAN WILLIGENBURG

*Department of Philosophy,
Erasmus University Rotterdam, The Netherlands*



KLUWER ACADEMIC PUBLISHERS

DORDRECHT / BOSTON / LONDON

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 0-7923-5304-8

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

The publication of this book was made possible by financial support from
The Netherlands Organisation for Scientific Research (NWO)

Printed on acid-free paper

All Rights Reserved

© 1998 Kluwer Academic Publishers and copyright holders
as specified on appropriate pages within.

No part of the material protected by this copyright notice may be reproduced or
utilized in any form or by any means, electronic or mechanical,
including photocopying, recording or by any information storage and
retrieval system, without written permission from the copyright owner.

Printed in the Netherlands

Contents

Wibren van der Burg and Theo van Willigenburg
Preface vii

Wibren van der Burg and Theo van Willigenburg
Introduction 1

Part I. The Elements of Reflective Equilibrium

Marcel Verweij
Moral Principles: Authoritative Norms or Flexible Guidelines? 29

Theo van Willigenburg
Morally Relevant Facts: Particularism and Intuitionist Rationality 41

Anton Vedder
Considered Judgements: Meaning, Community and Tradition 55

Ton van den Beld
Background Theories and Religious Beliefs: Their Role and Relation
in Reflective Equilibrium 73

Wibren van der Burg
Ideals and Ideal Theory: The Problem of Methodological Conservatism 89

Part II. The Methodology of Reflective Equilibrium

Eberhard Herrmann
Rationality, Warrant and Reflective Equilibrium 103

Marian Verkerk
The Thinker and the Thinking Process: A Feminist Perspective
on the Moral Faculty 115

Bo Petersson
Wide Reflective Equilibrium and the Justification of Moral Theory 127

Inez de Beaufort
Your Intuition or Mine? 135

Part III. The Scope of Reflective Equilibrium: Fundamental Views of Life and Religion

Vincent Brümmer
Religious Models and Moral Principles 145

<i>Carl Reinhold Bråkenhielm</i> Reflective Equilibrium and Cosmological Claims	157
<i>Anders Jeffner</i> World Views, Value Change and Reflective Equilibrium	171
Part IV. The Purposes of Reflective Equilibrium: Modelling Public Debate and Dialogue	
<i>Jan Vorstenbosch</i> Reflective Equilibrium and Public Debate: How to Cast the Public's Web of Beliefs Broadly Enough	177
<i>Frans W.A. Brom</i> Developing Public Morality: Between Practical Agreement and Intersubjective Reflective Equilibrium	191
<i>Jos Kessels</i> The Case of the Shared Values: An Example of a Socratic Dialogue	203
Part V. Case Studies in Bioethics	
<i>Egbert Schrotten</i> The 'Herman Case': The Usefulness of the Wide Reflective Equilibrium Model for Ethics Committees	219
<i>Bart Rutgers</i> The Use of the Reflective Equilibrium Method in Normative Veterinary Ethics	231
<i>Göran Collste</i> Infanticide in Reflective Equilibrium	239
<i>Johannes J.M. van Delden and Ghislaine J.M.W. van Thiel</i> Reflective Equilibrium as a Normative-Empirical Model in Bioethics	251
Bibliography	261
List of Contributors	277
Name Index	281
Tabula Gratulatorum	285

Introduction

Wibren van der Burg and Theo van Willigenburg

1 Reflective equilibrium: An attractive approach to moral thinking

Since John Rawls, in *A Theory of Justice*, elaborated the method of reflective equilibrium, it has gained much attention and popularity. Rawls introduced the method as a way of developing a general theory of justice, i.e. as an argumentative method for developing and justifying theories and principles to evaluate institutions and practices (Rawls 1971, 1985, 1987, 1993). Subsequent work by others has shown that a reflective equilibrium approach may be successfully used for various other purposes as well. It offers a promising coherentist account of justification, but it also gives guidance to philosophical thinking in more specific, practical contexts (e.g., Daniels 1979a, 1988; Heeger 1990a, 1992a and b; Kymlicka 1993; Carens 1995). Reflective equilibrium has especially proved to be a useful method for tackling practical moral problems (Beauchamp and Childress 1994).

The basic idea behind reflective equilibrium is that, in developing and justifying moral theories and in seeking answers to practical moral problems, we bring to bear – in some ordered way – all kinds of moral and non-moral beliefs and theories. A reflective equilibrium process pays attention to our moral and non-moral beliefs at various reflective levels (particular intuitions, moral principles, abstract theories), and ‘tests’ various parts of our belief system by revising and refining beliefs at all levels. In a process of mutual adjustment, we seek coherence among the widest possible set of beliefs that are arguably relevant in establishing a moral theory, in selecting moral principles or in deciding a specific moral problem.

The method of reflective equilibrium is clearly coherentist in that it does not favour a specific type of belief (Lyons 1975; Daniels 1979a, 1980b and c; Raz 1982; Hanen 1983; Sayre-McCord 1985, 1996; DePaul 1986, 1993; Brink 1987; Nielsen 1993).¹ There are no axioms upon which other beliefs can be based. Critical scrutiny is reached by questioning the tenability and relevance of all types of beliefs, none of which is immune to revision. If a person’s judgement regarding what is to be done in a particular case runs counter to what

1 However, some authors, e.g. Ebertz (1993), have argued that Rawls’s version of reflective equilibrium is less coherentist than he claims it to be; Ebertz calls him a moderate foundationalist.

is to be done in that case on the basis of a general principle he is committed to, then, according to the method of reflective equilibrium, it is an open question whether the judgement or the principle should be retained. So, in Rawls's famous words: 'Justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view.' (Rawls 1971, 21 and 579) When we talk about reflective equilibrium, we can refer both to a thinking or an argumentation process and to the product resulting from the process (Wellman 1971; Norman 1998). Product and process can never be separated because they presuppose each other. The process is one of a mutual adjustment of beliefs with the aim of attaining a full reflective equilibrium, a coherent set of beliefs, often in the form of a normative theory (Jamieson 1991). Only by looking at the process can we find out whether this equilibrium deserves the term reflective. If we merely look at the product in isolation, we can judge whether it is coherent, but a reflective equilibrium method requires more than simply coherence.

This basic idea of a process of mutual support and mutual adjustment between various beliefs is quite natural and intuitively appealing. Such an approach roughly approximates the way in which many of us tend to think when we are dealing with practical moral problems (Dworkin 1978; Jamieson 1991). According to Norman (1998), the method amounts to little more than a codification of common sense. It is not surprising, therefore, that the reflective equilibrium method is especially popular in various fields of applied ethics. Some authors have even argued that there is no real alternative (Grice 1978; Rääkkä 1996-97). Moreover, the idea of going back and forth between more general convictions, such as principles and laws, and more concrete elements, such as considered judgements, judicial decisions or empirical data, is also common in other practices like law and science (Dworkin 1978, 160 ff.; Hanen 1983).

2 Reflective equilibrium: A controversial approach

The reflective equilibrium method seems to be based, then, on a very natural idea, practised in many contexts of reflection and inquiry. But this does not mean that the idea is unproblematic. On the contrary, reflective equilibrium is a quite controversial method. Norman Daniels characterised the method as a 'process of bringing to bear the *broadest evidence* and *critical scrutiny*' (Daniels 1996, 2-3; our emphasis). It is precisely these two characteristics of (1) 'broadness' in evidence and (2) critical scrutiny by way of mutual refinement and revision of beliefs which have raised serious criticism and controversy.

We will start with the second point: the method for critical scrutiny. Opponents of reflective equilibrium argue that a coherentist method will never

result in sufficient justification for our moral judgements or theories (Hare 1971, 1996; Lyons 1975; Little 1984; Sencerz 1986; Haslett 1987; Timmons 1987; Stein 1994). And if we cannot claim sufficient justification for our moral judgements, how can we be justified to act on them? Critics particularly point to the risks of relativism and subjectivism. Some suggest that reflective equilibrium brings us nothing more than a neat systematisation of our prejudices (Hare 1971; Brandt 1979; Blackburn 1993).

The reply to such criticisms is quite straightforward (Nielsen 1977, 1982a and c, 1988, 1993; Swanton 1991). Perhaps the most annoying thing about the reflective equilibrium method is that it even takes pride in its anti-foundationalist way of justifying judgements, principles and theories. It does not even *try* to seek a more certain basis for normative statements, in contrast to the vast majority of philosophical (and theological) theories about morality that claim to find a basis in some ultimate principle or metaphysical claim, in divine commands or in the logic or pragmatic of normative language.² Neither does reflective equilibrium find a locus of certainty in particular or general moral intuitions revealing moral truth or in the wise man's appreciation of the moral character of a problem situation. The reflective equilibrium method tells us that it is the process of broad reflection as a whole that is important, rather than one of the epistemic elements in this process. Moreover, reflective equilibrium does not pretend to produce certainty; it aims at most at some weaker form of warrant. Any equilibrium attained is only a provisional equilibrium; it is always open to revision (Nielsen 1994b, 112).

This openness to revision is an important advantage of reflective equilibrium in dealing with criticisms and alternatives. It means that these should be seen as challenges, as a critical input into the reflective equilibrium process. Some critics have argued that it is not clear why having reached such a provisional equilibrium should make us think that we are any closer to knowing what is morally right or true (Hare 1971; Haslett 1987; Gibbard 1995). The natural answer of the reflective equilibrium theorist to such a criticism will be that the method invites us to throw in any consideration that makes one think that a provisional equilibrium does not bring us closer to moral truth, and use this consideration to test and adjust the equilibrium. This will at least lead to an improved equilibrium set of beliefs (Daniels 1979; Nielsen 1988).

Reflective equilibrium can incorporate whatever partial truth – in the form of criticisms, arguments, sets of beliefs – that traditional foundationalist moral approaches would be eager to bring to the fore. It will, therefore, often present a kind of middle ground between various competing alternatives (Daniels 1996,

2 As Nielsen phrases it (1994b, 90), it 'travels metaphysically and epistemologically light'. Cf. Norman (1998).

339). It may use Kantian arguments or utilitarian principles to criticise and to strengthen its analysis (Rawls 1980; Korsgaard 1996a). It may incorporate the insights of modern casuistry and virtue ethics (Hampshire 1983; Jonsen and Toulmin 1988; Van Willigenburg 1998b). In theory, it can include every line of argumentation and every type of belief; it need not exclude anything.

This strength, however, is also a major weakness, and here we find the second major criticism. If the method aims at being all-inclusive (aiming at bringing to bear ‘the broadest evidence’), it will soon be unworkable. Perhaps a Hercules or an archangel could combine all possible elements into one reflective equilibrium process and test and adjust them until some equilibrium is reached. For ordinary humans, this is impossible. It seems, therefore, that reflective equilibrium thinking requires some form of selectivity (Norman 1998). Though reflective equilibrium as a general approach is all-inclusive, every attempt to use it as a practical method must be selective. Choices must be made as to the types of beliefs, arguments and methodological criteria to be included, and how. As soon as we make such choices, we exclude other elements and criteria, and are thus vulnerable to criticisms of bias.

This need not be a problem that frustrates the method. Different versions of reflective equilibrium – all more or less biased in some sense – may serve different purposes; it depends on the purpose which biases are acceptable and which are not. For example, a methodological instrument like the Rawlsian veil of ignorance excludes knowledge of concrete facts and could thus lead to unrealistic outcomes (Miller 1992; Klosko 1993). As long as we focus on abstract principles for a political system, however, the resulting biases may remain tolerable. In other contexts, for example decision making in medical ethics, it would be absurd to exclude knowledge of concrete facts. This shows that a version of reflective equilibrium using an argumentative device such as the veil of ignorance can be acceptable for some specific purposes and completely unacceptable for other purposes. It appears, then, that we should not think of one all-purpose method of reflective equilibrium, but that we have to develop various versions of the method to suit different purposes (cf. Holmgren 1989, 60).

3 Two related perspectives on reflective equilibrium

From the previous section two related perspectives on reflective equilibrium and its history can be derived. Someone who is primarily interested in issues of moral epistemology will take the two major theoretical criticisms mentioned above as a starting point and analyse how various authors have tried to respond

to them. On the other hand, someone who is primarily interested in a practical use of the method will analyse the different versions of the method and the way in which they have been elaborated in order to serve specific purposes. In this introduction, and in this book at large, both perspectives, the epistemological and the pragmatic, are taken into account.

It will become clear that the epistemological and pragmatic questions concerning the wide reflective equilibrium method are intertwined. This has to do with the nature of the method, which claims to show us a systematic manner of conducting our moral inquiries. Reflective equilibrium is meant to structure our moral arguments when we try to assess acts and policies or to develop and justify theories and principles for evaluating institutions and practices. Justification along the coherentist lines as sketched by the reflective equilibrium method should be understood as a process of persuading others (and ourselves), whereby 'others' may include any reasonable challenger. The method aims at providing guidance in concrete moral decision making or in building theories that can be taken as reliable guides. This goal partly explains the specific epistemological characteristics of the method, like the important role it assigns to moral intuitions or considered judgements (cf. Rawls 1985, 228 f.). The best way to convince reasonable challengers is to show that the results of reflective equilibrium match their considered judgements, or to show – if the results do not match the initial judgements of the challengers – that, because of other beliefs they endorse, they would come to find these results convincing. A coherentist approach like wide reflective equilibrium best suits the way in which moral inquiry proceeds in its mundane and day-to-day form.

This strongly pragmatic focus raises epistemological doubts and questions which cannot be ignored, however. What reasons do we have to think that a method like this will really deliver on its promise of progress and convergence in moral argument and theory? And if there is convergence of belief, what basis do we have for thinking that this has brought us closer to what we were looking for: a reliable moral assessment of acts, policies and institutions? In the next section, we will take up these epistemological questions. Surprisingly enough, it will appear that – in the end – these questions can only be fully answered by choosing again a more pragmatic perspective. If wide reflective equilibrium is best conceived as a model of persuasive argumentation – justification being the condition of standing up to all reasonable challenges – then the epistemic value of the result of reflective equilibrium thinking will strongly depend on the context and goals of argumentation.

4 Reflective equilibrium method and credibility

In 1971, Rawls used the reflective equilibrium method to develop a theory of justice that was to describe and systematise our 'sense of justice'. The theory he aimed to construct was to contain a core set of principles, which together would build a 'conception of justice' that characterises our moral sensibility with regard to questions of justice – a sensibility which is part of our 'moral capacity'. Given this idea of the theory of justice looked for, a number of rival theories were tested by applying their principles and seeing whether the results match our 'considered judgements'. The latter are those 'judgments in which our moral capacities are most likely to be displayed without distortion' because they are given 'under conditions favorable for deliberation and judgment in general' (Rawls 1971, 48). There is a certain ambiguity here. On the one hand, Rawls insisted that considered judgements are not immune to criticism and testing, which means that our moral capacity may itself alter under the influence of the theory and its principles.³ On the other hand, reflective equilibrium was not presented as a full-blown coherence method, but was modelled on scientific method, which describes theory construction as proceeding from observation reports (English 1978; Nielsen 1982a; Gutting 1982).⁴ Just like in science, theoretical conjectures are tested against the 'facts' as they appear in observation reports. '[T]here is a definite if limited class of facts against which conjectured principles can be checked, namely our considered judgments in reflective equilibrium' (Rawls 1971, 51).

This analogy with scientific method is even more explicit in Rawls's early article 'Outline of a Decision Procedure in Ethics' (1951). In this article, he articulated a method for validating and invalidating proposed moral principles and decisions made on the basis of them in concrete cases. The suggested decision procedure consisted of three stages: (1) a procedure for isolating a set of moral 'data', (2) a procedure for discovering and formulating a set of principles which would satisfactorily explicate these data, and (3) a discussion of the reasons one might have for accepting these principles as rationally justifiable. The main part of the 1951 article consisted of an elaboration of criteria on the basis of which a judgement may be called 'considered' and, thereby, function as one of the moral 'data' against which principles are tested. Rawls not only formulated various conditions on the nature and emergence of

3 This line of argument corresponds with what Dworkin (1978) called a 'constructive' model of reflective equilibrium, which he regarded as the best interpretation of Rawls (1971).

4 This line of argument corresponds with Dworkin's 'natural' model of reflective equilibrium, which aims to describe an objective moral reality (Dworkin 1978).

considered judgements, but also defined a limited set of 'competent judges' (characterised by a certain requisite degree of intelligence, knowledge, rationality, sympathy and 'moral insight'), whose considered judgements are the only ones which count as reliable data for decision making and theory construction.

In *A Theory of Justice* (1971), this idea of a limited set of competent judges is abandoned. An appeal can be made to the considered judgements of 'bien-pensants' generally, and Rawls even indicates that 'for the purpose of this book, the views of the reader and the author are the only ones that count' (Rawls 1971, 50). The criteria for calling a judgement considered have also changed and have become more general: 'Considered judgments are simply those rendered under conditions favorable to the exercise of the sense of justice, and therefore in circumstances where the common excuses and explanations for making a mistake do not obtain' (*Id.*, 47).

The reason for this shift is that in *A Theory of Justice* considered convictions are touchstones in a more modest sense than in Rawls's earlier description of his method. Rawls now more openly embraces a coherentist epistemology, according to which the data for which the theory has to account themselves may be adjusted, rectified and even rejected because of some convincing principle (Delaney 1977). There is no one-way accommodation but 'mutual adjustment', a concept which Rawls adapted from Goodman's philosophical account of inductive inference (Goodman 1955 – see Rawls 1971, 20 n. 7). Rawls (1971, 579 n. 33) also refers to the conception of justification found in the work of Quine (1960), where Quine stresses the role of consensus in the definition of 'observation'.

Even so, Rawls still does not seem to endorse a radical coherentism. He compares the reflective equilibrium method with Chomsky's attempt to formulate principles which describe 'the sense of grammaticalness that we have for the sentences of our native language' (Rawls 1971, 20). Analogously, moral philosophy aims at the formulation of principles which account for a person's sense of justice (Goodman 1955, 65 f.; Daniels 1980a; Singer 1986). This 'sense of justice' will be subject to scrutiny because the judgements based on it will be confronted with principles building a conception of justice, but it is hardly imaginable that this scrutiny will lead to radical alterations if the analogy with grammar holds.

Rawls's seemingly half-hearted coherentism has led to a long-standing debate about the status and nature of considered moral judgements (Carr 1975; Nielsen 1977 and 1982c; Grice 1978; Noble 1979; Sencerz 1986; Barry 1989a, 271 f.; Nelson 1990; Audi 1993; Ebertz 1993; Kekes 1993). Rawls seems to award a special epistemic status to considered moral judgements in order to give the

process of reflective equilibrium some reliable starting points. This may be an effective way of answering those critics who think that pure coherence arguments will not help us in making progress in moral thinking as they simply move us in circles around our current considered opinions and the principles they supposedly manifest.

But what could give considered moral judgements their epistemic credentials? There is a decisive difference here compared with scientific observation reports. We can explain why observation reports are generally reliable and these (causal) explanations do not completely rely on the credibility of the body of theoretical beliefs which provide their explanation (Dancy 1984; see also Dancy 1985 and Haslett 1987). There is, however, no evidence for the reliability of considered judgements independent of the principles and other theoretical elements they are supposed to support and manifest (Sencerz 1986). In Richard Brandt's terms: the initial *credence* level of considered moral judgements, i.e. the level to which we are committed to the beliefs involved in these judgements, does not tell us anything about their *credibility* (Brandt 1959, 1979, 1989, 1996). The fact that we trust our well-considered judgements does not add any evidence to their trustworthiness. Nothing prevents our considered judgements from expressing just the arbitrary commitments and sentiments of a prejudiced viewpoint (Singer 1974; Blackburn 1993). Increasing the credence level of our considered judgements by making them cohere with principles that provide generalisation and systematisation does not move us one inch closer to credibility. And there is nothing in the further process of mutually adjusting considered judgements and principles that could give us reason for thinking that we had moved to a more credible viewpoint instead of having refined and confirmed our initial prejudiced outlook.

In answer to this no-credibility objection, Norman Daniels developed the idea – already manifest in Rawls's work – of a *wide reflective equilibrium*. The idea is to seek coherence among the *widest* set of moral and non-moral beliefs by revising and refining them *at all levels* (Daniels 1979a; for a critical view, see Holmgren 1989). Reflection is not limited to principles and considered judgements. Reflection also has to include relevant background theories like a theory of the person or a general social theory, specific thinking formats like the original position, and specific concepts like the concept of a social contract. These background theories, thinking formats and concepts may provide evidence for the credibility of a set of moral principles and the moral conception they embody.

Explicitly involving background theories in the method of reflective equilibrium provides an answer to the often heard criticism of Rawls's theory of justice that it implicitly presupposes some theory of the person (e.g. a theory which emphasises the separateness of free and equal persons), some theory about

the role of justice in society (including the ideal of a well-ordered society) and various social theories necessary to back up Rawls's assumption about the feasibility and stability of a system based on procedural fairness. (See various contributions in Daniels 1975; Hampton 1989; Raz 1990; see also Sandel 1982; Rawls 1985 and 1987; Barry 1989a, 320f.; Nielsen 1994b). Wide reflective equilibrium explicitly includes all these background theories in the process of mutual reflective support and adjustment and, thus, expands the kinds of considerations that count as evidence for or against our moral views at all levels of generality. By paying attention to all kinds of moral and non-moral considerations that bear on a moral issue, it presents a model of moral thinking which, in its idealised form, may provide an answer to the criticism that coherence of beliefs does not lead to credibility. The moral thinker using the wide reflective equilibrium method has a simple question to put to the sceptic who wonders whether the established wide equilibrium of considerations does not just build a coherent set of prejudices: What consideration makes you think that this is just an amalgam of cooked-up prejudices? What consideration gives you reason to be suspicious? And when the sceptic comes up with this consideration, the moral thinker will invite her to throw these critical considerations back into the ring and seek a new wide reflective equilibrium. It seems plausible to suppose that enriching wide reflective equilibrium in this way contributes to the credibility of the judgements contained in it (Daniels 1979a).

What could the sceptic say to this? There are at least two types of criticism that the sceptic could bring which are worth mentioning here because they have inspired the further development of the wide reflective equilibrium approach.

The first type of criticism focuses on the conservative tendency inherent in any method of thinking that starts with our actual considered moral judgements.⁵ The more epistemic value is bestowed on the considerations which are taken as starting points of reflective equilibrium thinking, the greater is the danger of including biases which can never be traced nor erased. Widening the circle of considerations to be included in reflective equilibrium does not bring us closer to moral truth if the point where we actually start drawing the circle is simply given with our initial judgements, even if these are formed in optimal epistemic circumstances.

In answer to this line of criticism, Michael DePaul (1993) proposed a *radical conception of reflective equilibrium*, which allows initial beliefs and degrees of beliefs to be altered in ways that go beyond what is required to make

5 Singer (1974); Brandt (1979); Aronovitch (1996); cf. also Cladis (1994); Walker (1995); Rääkkä (1996-97, 185). Nielsen 1991 argues that reflective equilibrium is historicist and contextualist but not relativist.

a person's initial belief system coherent. In wide reflective equilibrium as conceived by Norman Daniels, the raw materials (initial judgements) play an important role in determining the nature of the final coherence set because refinement and revision are only called for in order to resolve conflicts where these initial beliefs are involved or to extend the system of beliefs where the extension is always extension beyond the initial beliefs. Instead, a method of radical reflective equilibrium may result in a coherent set of beliefs which is inconsistent with any of the initial beliefs with which the thinking process started. The radical method may lead, so to speak, to radical conversions in one's thinking. Initial beliefs may start the process of reflection, but they will in no sense determine the direction this reflection may follow. Very different starting points may lead to the same equilibrium point of thinking, but the same starting point may also bring different people to very different outlooks on an issue. There is, therefore, no need to be worried about the credibility of the initial moral judgements.

DePaul has an interesting suggestion about how to promote such a radical revision process. He makes clear that the highly intellectualist focus on moral judgement, construction of arguments and development of theories of standard reflective equilibrium approaches is one-sided. It leaves out important questions about the character and development of the inquirer's *abilities* and *faculties* for making judgements, constructing arguments and building theories. Typically, a person can acquire the ability to make relevant discriminations in judging and arguing only after a considerable amount of experience and training. Philosophical deliberation along the lines of the reflective equilibrium method should not only be thought of as affecting our beliefs and arguments, but should also be expected to cause 'changes in a person's judgmental faculties, so that these faculties no longer function in the same way, yielding the same beliefs and theories, as they previously did' (DePaul 1993, 211). This means that reflective equilibrium thinking may demand the expansion of one's range of experiences – be they 'real-life' or vicarious experiences through, e.g., literature and theatre – in order to test, refine and possibly revise one's judgmental faculties.

DePaul thus draws our attention to an important point which, in the Rawlsian original position, is deliberately ignored as much as possible, but which should not be neglected: the character and experience of the person who goes through the reflective equilibrium process. Moreover, he gives an interesting suggestion about how to counterbalance possible conservative tendencies. The question remains, however, whether this will indeed bring us closer to moral beliefs which are true, or at least warranted.

The second type of criticism of the wide reflective equilibrium approach points to some strong reasons to limit our expectations with regard to the results attained by using a method of philosophical thinking like wide reflective equilibrium (Baier 1985; Williams 1985; Clarke 1987; Raz 1990; Gibbard 1995). Even the most reasonable and most experienced people are likely to develop quite varied answers to the complex moral issues that humans seek answers to. This pluralism is a pervasive and deep fact; in his recent work, Rawls admits that, in *A Theory of Justice*, he was insufficiently sensitive to it (Rawls 1993, 54 ff.). Because of this pluralism, it is unrealistic to expect philosophical reflection – e.g. along the lines of reflective equilibrium – to move everyone who can think clearly and rationally, and who bases his judgement on a rich and uncorrupted experience, to convergence on one moral outlook *vis-à-vis* certain matters. People may, for instance, have very different views about the priority of a value like justice. For a religious person, a relationship with God may be much more important than distributive justice. Others may think that being part of a caring community is much more important than a just society. Why would these people have to sacrifice their important moral concerns in order to abide by the principles of justice? Is the wide reflective equilibrium theorist who thinks that, in the end, the views of people will converge on some overall coherent set of moral considerations, not blatantly naive?

In answer to this type of criticism, John Rawls tries to show that a *political* conception of justice may be reflectively endorsed even if there is no general convergence on a wide reflective equilibrium with comprehensive commitments that support this political conception (Rawls 1985, 1989, 1993). Philosophical, moral and religious controversies should be left aside as much as possible in ‘a method of avoidance’ (Rawls 1985, 283; Cladis 1994).⁶ Daniels calls this a ‘*political reflective equilibrium*’ in which, next to philosophical reflection, a major role is assigned to shared institutions. These institutions make, in the course of a shared history, groups with different comprehensive views accommodate themselves to justice as fairness (Daniels 1996, 160). These shared institutions have a distinctive political character. They make it possible to create a public or political realm of reasoning about matters of justice. In this realm, people share the key ideas of justice as fairness, like the idea that citizens are free and equal – even if they support these ideas for very different reasons. Different comprehensive views may lead to different justifications for the principles, reasons and ways of argumentation that are internal to the political conception of justice; yet, there is an overlap in the outcomes. They will even

6 Dworkin’s constructivist interpretation of reflective equilibrium and the justification he gives for this interpretation in terms of a public standpoint already suggested some of the ideas elaborated in Rawls (1993). Cf. Dworkin (1978, 163).

imply different views on where to draw the line between the political and the non-political. Despite these differences, Rawls assumes that there is enough 'overlapping consensus' to create a commonly shared 'political reflective equilibrium' (Rawls 1987, 1989, 1993; see also Nielsen 1991 and 1994b; Cohen 1994; Greenawalt 1994; Barry 1995).⁷

5 The pragmatics of reflective equilibrium

The various elaborations on reflective equilibrium – whether they constitute wide, radical or political conceptions – have certainly led to improvements. Nevertheless, they do not provide knock-down arguments to prove that it is the best method of moral reasoning, let alone that it leads to moral truth (if there is such a thing as moral truth). We can never prove a priori that it is so; we can only demonstrate that it does well in comparison with other methods. This brings us to the pragmatics of reflective equilibrium because we cannot compare argumentation methods in general, only in connection with the purposes they are meant to serve.

A reflective equilibrium process may deliberately include or exclude certain types of considerations, depending on the issues at stake and the purpose aimed at. The way in which methodological requirements, such as the ideal of impartiality, are interpreted differs similarly. So there may be different types of reflective equilibrium. This pragmatic selectivity in the light of specific purposes can be an adequate response to the danger of all-inclusiveness, which could make a reflective equilibrium method unworkable. Moreover, it is also a good method to counter some items of epistemological criticism mentioned earlier. Rather than attempting to construct one general method, one can carefully analyse which biases are likely to occur under specific circumstances and design methodological devices or include specific elements in the process to counter these biases most effectively.

We can structure the possible varieties of reflective equilibrium methods by asking three questions:

1. What are the purposes the method is meant to serve?
2. What are the considerations used in the process?
3. What are the methodological requirements to be observed in the process?

7 Klosko (1993, 353), however, argues that the idea of an overlapping consensus on liberal ideas fails on empirical grounds, at least for concrete issues: 'The unavoidable conclusion of decades of empirical studies, then, is that significant majorities of Americans, like citizens of other liberal societies, do not uphold basic liberties in the strong sense Rawls requires.' For a similar criticism see Walker (1995, 102).

The variation in purposes of reflective equilibrium and thus in the method itself can, to a certain extent, also be found in the work of John Rawls. In 1951, he regarded the method as a decision procedure and its product as a set of moral principles. The methodological requirements focus on the criteria for considered judgements and on the qualification of 'competent judges'. The considerations used in the process were restricted to considered judgements ('moral data') and moral principles.

In 1971 the method was used for theory construction: its aim was, to be more specific, to arrive at a normative theory (and especially the basic principles) for the basic structure of an almost ideal (nearly just) society. Methodological devices such as the original position and the veil of ignorance were introduced and controversial religious views and knowledge of concrete facts were excluded from the argumentation process (cf. Daniels 1975).

In 1993 there was another change of purpose because the aim no longer was a theory for an ideal society but a normative political theory for a modern, deeply pluralist society. In the method, the idea of an overlapping consensus was given a central role, but controversial religious views were still excluded from the reflective equilibrium process (cf. Barry 1995).

Despite the variation, the primary purpose of the Rawlsian versions of reflective equilibrium is theory construction and the construction of moral or political principles; they tend to focus on the more abstract theoretical levels (Barry 1989a). Even the 1951 version only results in moral principles which still have to be interpreted and applied to concrete problems, whereas the other versions need various steps of transformation to be useful for practical problems. Therefore, they are less adequate for the purpose of concrete problem solving; we should look elsewhere for versions that may be useful in applied ethics.

To get a new start, we should return to the minimal idea of reflective equilibrium (as in the 1951 article), which has been the basis for much work in applied ethics. Two elements are at the core of every version of reflective equilibrium: considered moral judgements and general moral principles or general beliefs. A method which consists of no more than these two elements (and some methodological requirements) is usually called a narrow reflective equilibrium (Daniels 1979a). Some authors have, for their purposes, been satisfied with this minimal version. It is, e.g., used by Glover (1977, 26 f.), who speaks about an interplay between 'responses' and 'general beliefs'. In the early versions of their textbook on biomedical ethics, Beauchamp and Childress (e.g., 1983, 12) similarly described their procedure as a dialectical reasoning between principles and judgements about cases.⁸

8 In the fourth edition of their book (1994, 20-37), however, they explicitly subscribe to a very wide version of reflective equilibrium.

Most authors in applied ethics have nevertheless gone beyond the narrow reflective equilibrium and added or refined some elements. If one takes considered judgements as having rather specific ideas and particular situations or cases as their object, and principles as the more general normative propositions, one could introduce 'morally relevant facts' as a separate element in reflective equilibrium. This will be particularly useful if one uses reflective equilibrium as a method for analysing concrete moral problems. For, in order to judge a situation, we must distinguish the morally salient features of that situation, i.e. we must point out some natural properties of the case as a reason to evaluate the situation in one way or another. Whether one thinks it possible or even necessary to introduce morally relevant facts as distinct elements into the reflective equilibrium process, seems to depend partly on whether one believes that moral salience will 'simply' appear if one is sensitive to the situation (morally salient facts 'leaping' to the eye) or whether one believes that salience is brought in by more general moral considerations, that is, considerations that can be formulated independent of the concrete case.

In the first line of thought, Robert Heeger and Theo van Willigenburg have developed a 'network model' for dealing with concrete moral problems (Heeger 1990a; Van Willigenburg and Heeger 1991). This network model consists of three basic elements: moral intuitions, moral principles and morally relevant facts. They also suggest various methodological requirements regarding these elements (such as the principles' universality and simplicity) and regarding the ethical practitioner (such as an open mind and moral sensibility). This is a very simple model, basically the narrow reflective equilibrium to which morally relevant facts have been added. It is important to notice, however, that the facts are not merely regarded as passive objects to which moral judgements or principles have to be applied, but also as a factor in the equilibrium process itself. The (Rossian) idea is that, being confronted with a morally problematic situation and having some idea of relevant prima-facie duties, people naturally form intuitions about the overall moral story to be told about the situation, which includes some initial intuitive view on what is morally required. There is thus a genuine interplay between facts, principles and intuitions.

This network model has subsequently been elaborated and refined in a number of doctoral theses and other publications. It can be argued that methods such as the network model (or narrow reflective equilibrium) are most productive and least vulnerable to criticism when applied to fields in which there is already a relatively stable (though still partly incomplete) moral consensus both at the level of moral principles and at the level of concrete issues (Norman 1998). In such contexts it may be very useful for problem solving. (For an example of such an application of the network model, see Rutgers 1993.) However, in fields in which there is less consensus, in which moral opinions are in rapid

change and moral philosophers are still searching for defensible formulations of basic principles and values, such a model is less adequate. When we need to develop moral views about radically new developments such as modern biotechnology, we need to broaden the scope and explicitly refer to background theories. It may even be unavoidable to include controversial background views such as religious views in order to deal with these issues (cf. Heeger 1997).

The network model should perhaps be further broadened if we want to deal with practical issues which have broader theoretical implications or if we want to develop 'modest theories' or guidelines. One change may be necessary if we want to construct, e.g., a practically useful democratic theory for the real society we live in; it may then be useful to add the basic values or ideals that are implicit in the public culture of our society. We thus get an 'extended network model' in which five elements are to be included: moral intuitions, moral principles, morally relevant facts, background theories and moral ideals (Van der Burg 1991; Brom 1997).

Another change may be needed if we use reflective equilibrium to develop a set of guidelines for concrete decision making. Hans van Delden, e.g., uses it to construct a set of guidelines on do-not-resuscitate decisions (Van Delden 1993). He makes an interesting change by not starting from the intuitions and principles as formulated by the philosopher but from the intuitions and principles of the practitioners; he incorporates the results of empirical research on moral opinions into the model.⁹ By way of this strategy, he increases the chances that his recommendations will be recognised by doctors and nurses as a critical reconstruction of their own views and, thus, that they will be accepted. Furthermore, as his purpose is effectiveness of his recommendations in practice, he suggests that these guidelines be supplemented with a virtue theory. Again, the purpose (in this case the development of a practically effective set of guidelines) determines both the elements and the method of reflective equilibrium and, in a sense, even requires going beyond the equilibrium method as such.

Thus, different versions of the network model may be used for different purposes in the context of applied ethics. Yet other varieties of reflective equilibrium could be useful if the focus is more on intersubjective discussions and public debates. Usually, reflective equilibrium is merely a method for

9 The idea that empirical research on moral views is relevant to reflective equilibrium methods has been suggested by various authors (Ebertz 1993, 208 ff.; Miller 1994), yet attempts to integrate them are still scarce. Another example is Klosko (1993).

intrapersonal argument.¹⁰ It may, nevertheless, be used as a model for interpersonal argumentation, as a real-life framework for structuring discussions. For example, we could structure a discussion in ethics committees along the lines of a reflective equilibrium process. It may also be regarded as a framework for structuring political debates in parliament or even in society at large. Each of these purposes will bring the need to modify the model, either by adding or ignoring certain elements in the process, or by constructing new methodological requirements. For instance, if we want to use reflective equilibrium as a model for committee discussion, we must see how the idea of impartiality can be guaranteed in a committee, both by the process of selecting members and by structuring the discussion and the attitude expected from the members.

Reflective equilibrium, thus, need not only be a method for philosophers – it may even be more attractive and effective for the general public. It may also be a method for moral education and, in connection with this, a model for moral development. The student can be confronted with concrete moral problems and be asked to solve them, starting with her moral intuitions, reflecting on them, trying to find principles that do justice to them, and so on. (In our experience, this basic structure proves to be very useful in teaching ethics to students in the various biomedical departments.)

There are even other, quite different purposes for which a version of the reflective equilibrium method may be helpful. We need not only think of reflective equilibrium as a normative model; it may also be useful as a descriptive model. This may be so at the level of the moral development of one person, but also at the level of society or societal subgroups. If it is a good method in normative argument, the hypothesis is warranted that at least sometimes moral developments in society follow the lines of a reflective equilibrium process (cf. Brom 1997).

We may conclude that there is a broad variety of purposes for which reflective equilibrium methods may be used and, corresponding with this variety, there are many different versions. A systematic comparison between these versions could be a further step in the development of reflective equilibrium. For example, detailed empirical studies of the structure of public debates may be helpful to test and refine the method in normative contexts. Seeing how a method works in the context of theory construction and which biases result from

10 Of course, Rawls (1971) presents his theory as one in which various persons convene to decide on principles of justice, but many critics have convincingly argued that, in fact, there is only one – imagined – person doing the reasoning. Aronovitch (1996, 401) argues that is the ‘typical person’ rather than a special person.

specific methodological requirements may make us sensitive for similar biases in models developed for problem solving.

A further refinement of reflective equilibrium methods may also be inspired by studies in other disciplines such as science and grammar. One practice where reasoning can be structured adequately along the lines of reflective equilibrium is law (Dworkin 1978; Hanen 1983; Van der Burg 1991 and 1993; for a critical view see Sunstein 1995). Haslett (1987, 309) has even argued that this method is more promising in legal than in moral reasoning because in case law and statutes, law has at least some elements with initial credibility. Specific characteristics of the versions in law may also be of heuristic value for a refinement of the method in ethics. For example, the virtue of integrity, which Ronald Dworkin has suggested is a central ideal in law and which may be a fundamental reason for striving towards coherence, could also be central to reflective equilibrium in (political) ethics. A comparison between how integrity functions in law, in morality and in politics may therefore result in a better understanding of the role of coherence in political and moral argument (Dworkin 1986; Van der Burg 1993).¹¹

6 The contributions to this volume

The contributions in Part I, *The Elements of Reflective Equilibrium*, illuminate and assess the nature of the various types of considerations or belief 'elements' which play a role in wide reflective equilibrium thinking. Marcel Verweij discusses the nature of *moral principles* by asking how principles can retain their character as more or less stringent and durable moral requirements – the cornerstones of one's system of moral beliefs – and still be open to revision or even rejection in a reflective equilibrium thinking process. Verweij argues that if reflective equilibrium is employed as a method for theory formation, one may distinguish (i) initial principles, (ii) principles during the process of reflection and adjustment and (iii) principles that are the product of this process. Initial principles may have a specific appeal and, therefore, authority that cannot be reduced to the inquirer's commitment to his concrete judgements. (Principles are more than summaries thereof.) The authoritative appeal of initial principles may, in a sense, be retained during the reflective equilibrium process even if they are subject to revision and adjustment. The situation is different when reflective equilibrium thinking is employed to assess concrete moral problems. Principles seem to function, then, more as a kind of containers of a wide range

11 Dworkin (1978, 163 and 1986); cf. Shiner (1992, 229). Compare also a similar point in Glover (1977, 28), on scientific integrity as the basis for coherence.

of possibly relevant moral considerations. Still, some of these considerations will build the authoritative core of the principle (having, therefore, clear binding force), while others may be more in the penumbra of the meaning of the principle, requiring specification and thereby interpretation in the light of the problem at hand. But here too principles may retain some of their authoritative appeal and kernel even in a process of change and adjustment.

Theo van Willigenburg asks what makes certain *facts of a case* morally relevant, i.e. bearing moral weight in judging the case. What makes certain facts count as reasons for or against a particular moral judgement? He argues that case features may appear to have moral relevance prior to and independent of having generalisable considerations like moral principles or rules pointing out and explaining that relevance. The reason is that the relevance of certain facts may dramatically vary according to circumstances, because of the strong context dependence of what counts as morally relevant. Facts, therefore, can have moral salience *sui generis* in a certain case, which makes it impossible to fully 'catch' moral relevance by invoking general moral considerations. The recognition of *sui generis* relevant facts and the understanding of the way they play their reason-giving role in reflective equilibrium thinking demands a kind of non-inferential, intuitionist rationality (based on a particularist epistemology). Van Willigenburg claims that intuitionism may very well be combined with inferential reasoning involving the various generalisable considerations which play their role in reflective equilibrium thinking.

Reflective equilibrium is often understood as a method for testing our ordinary *considered moral judgements* by bringing them into critical interaction with more general moral and non-moral considerations. The main question to be posed against the method is what makes us think that adjusting and pruning our judgements following the method of reflective equilibrium will free us from moral bias and bigotry? Anton Vedder argues that the specific coherentist method of moral justification as exemplified by the reflective equilibrium method makes it suited – par excellence – to confer justificatory force on our considered moral judgements. Wide reflective equilibrium functions as an instrument for unravelling the web of moral and non-moral meaning embedded in the shared experience and knowledge of communities and traditions. By exploring and formulating the connections between various moral and non-moral notions, reflective equilibrium thinking enhances our understanding of the way our moral judgements are embedded in the contexts of meanings, which together build a life form of which we are part. This does not exclude controversy and difference within a tradition or community, but it makes clear how understanding and justification always are fundamentally context bound – a deep fact acknowledged by the method of reflective equilibrium thinking.

The contributions by Van Willigenburg and Vedder both stress the importance of a particular life form as the source of understanding moral notions and as the background of the way 'we' intuitively grasp the moral meaning of certain constellations of facts. Ton van den Beld analyses the role views of life – fundamental to a person's outlook and biography – should have in wide reflective equilibrium thinking. Are views of life to be included in the *set of background theories and considerations* which Daniels introduced into wide reflective equilibrium? Van den Beld shows that the metaphysical beliefs which a religious person may endorse, may have a direct influence on the way this person comes to weigh and balance various moral considerations in trying to reach an equilibrium of convictions with regard to a problematic issue. Having fundamental views of life 'at the background' of reflective equilibrium thinking does not mean, however, that the reflective equilibrium method would be structured in a more foundationalist way. Religious beliefs and other background theories and views may be subject to scrutiny in the course of wide reflective equilibrium thinking, just as other belief 'elements' like principles and considered moral judgements. Van den Beld argues that reflective equilibrium thinking aimed at reaching interpersonal moral agreement and convergence in moral enquiry – for instance in the political sphere – need not at the outset exclude 'controversial' fundamental views of life from the agenda. Rawls's idea of an 'overlapping consensus' of views leaves enough room for having religious beliefs play their role as part of the background beliefs in wide reflective equilibrium.

Important to any life form are values that are implicit and latent in the public and moral culture. Values which are thus grounded in reality but which are oriented towards the future are the *ideals* or normative ambitions of a group or person. Such ideals usually cannot be fully realised and they partly transcend contingent, historical formulations and implementations in terms of rules and principles. In his contribution, Wibren van der Burg clarifies the role and character of these ideals and argues that reflective equilibrium thinking should account for their normative force and their specific role as continuous sources of new critical input into the reflective equilibrium process. As ideals are, on the one hand, immanent in our culture or experience and, on the other hand, transcend this culture or experience, they may form a point of critical reflection because their endorsement provides access to an indefinite surplus of normative meaning and sensitiveness to unforeseen possibilities. Having ideals as reflective 'elements' in our thinking process may, thus, countervail the conservative tendency of the (narrow) reflective equilibrium method.

Part II on *The Methodology of Reflective Equilibrium* takes up some intricate epistemological problems involved in reflective equilibrium as a method of moral

enquiry. According to the method, it is rational to endorse a moral belief if this belief is sufficiently supported in a network of connected and coherent other beliefs. Usually, rationality is considered to be a person-related, epistemological criterion – what should be (ir)rational for you to know, given your set of beliefs, need not be identical to what should be (ir)rational for me to know. It is usually contended, therefore, that next to rationality we need another epistemological criterion, warrant, which introduces a general standard or procedure which may authorise the step from the evidential grounds to a claim, i.e. a procedure which explains why one may trust the conclusion reached. Warrant is truth conductive in the sense that it turns mere belief into knowledge. Eberhard Herrmann argues, however, that warrant in the truth-conductive sense is, just as rationality, an inevitably person-related criterion. Still – he says – it is possible to have a difference between rationality and warrant if one conceives – in a pragmatist way – truth as warranted assertability. In this view, knowing is not mentally mimicking what is ‘out there’, but is a constructive conceptual activity that guides our pragmatic interaction with the world. According to Herrmann, Putnam’s internal pragmatic realism – a realism concerning what humans can conceive of as ideal verification – gives us enough epistemological foothold to avoid the person relativism inherent in the concept of warrant. To warrant a statement is to consider our ‘epistemic circumstances’ as ideal as, within our human situation, we could hope to have them. This explains why, according to Herrmann, it is necessary to be in an optimal ‘condition’ not only cognitively but also emotionally, and have a broad range of so-called formative experiences which help us develop a sensitivity for good and evil, right and wrong. Therefore, Herrmann says, moral beliefs may become warranted through moral experience developed not only by philosophical reasoning but also by way of art and the way different views of life express themselves.

Marian Verkerk further develops Herrmann’s point about the importance of formative experience. She argues, however, that we must not only pay attention to the specific sensitivities (our moral ‘faculty’) which we need in morally judging situations and balancing considerations; we also have to pay attention to the social and political interpretative framework of our thinking. Feminist moral epistemology shows that whether someone has rational and warranted beliefs is not only a question of having adequate reasoning capacities, but also, and more importantly, a matter of her having a sense of her ‘self’ that would support a full sense of flourishing. Not only personal, but also social and political transformations and sensitivisations are necessary to allow ideologically oppressed people to acquire new interests, desires, beliefs about the world and understanding of their own position. Literature and other forms of art may be helpful in developing one’s capacities of moral understanding, but this should not distract our attention from the political aspects of defining and acquiring

an adequate sense of personal integrity, which is a basic condition for sound moral judgement.

Bo Petersson provides a critical analysis of the way coherence is employed as a justificatory criterion in wide reflective equilibrium thinking. There may be more ways to exclude incoherencies and to establish coherence of considerations, so there may be more than one equilibrium to be attained. Petersson shows that ideas of 'maximal coherence' and 'best fit' of beliefs are yet too unclear to provide sufficient guidance when it comes to a more systematic and detailed comparison between coherent sets of beliefs. It is, for instance, not clear what should count as decisive in judging 'maximal coherence'. Is strong mutual support through entailment relations of a small set of beliefs more important than supportive relations of a weaker kind (inductive or probabilistic or explanatory 'entailment') in a wide set of beliefs? Coherentist justification alone is not enough to attain warranted assertability. Petersson suggests that we further develop the main idea behind wide reflective equilibrium: our beliefs about the world, man, society and morality must fit together in a continuous and connected whole. This implies holism rather than coherentism, the difference being that (i) – if necessary – holism intends to endorse very different types of beliefs, and (ii) holism does not exclude the possibility that some kinds of belief have more weight than others without assuming foundationalism.

Not only is it possible that there is more than one reflective equilibrium point, i.e. more than one way to reach coherence, in moral thinking very different persons may be involved with very different intuitions. Which considered judgements need adjustment and correction, and which intuition is to be upheld against conflicting principles or theories? 'Your Intuition or Mine?' is the provoking title of Inez de Beaufort's contribution. She is deeply suspicious of the role of so-called 'yuk feelings' – strong feelings of 'this is wrong' – which seem to inspire many of the debates in medical ethics. Why would gut feelings have any authority in moral thinking? Why take intuitions seriously? Using the (up-to-now fictitious) example of male pregnancy, De Beaufort shows what the reasons could be for following and trying to understand our gut feelings about controversial issues (Are these feelings not grounded in the firmly held beliefs within a moral tradition?), and what reason there could be for mistrusting people's gut feelings (unjustified conservatism). The arguments De Beaufort brings forward for and against trusting our intuitions put some of the main features of reflective equilibrium thinking under serious criticism.

The contributions in Part III, *The Scope of Reflective Equilibrium: Fundamental Views of Life and Religion*, discuss the relations of moral commitments with religious beliefs and – correspondingly – the place of views of life in reflective equilibrium thinking. Vincent Brümmer argues that we could express our moral

commitments either in terms of moral principles or in terms of models and metaphors derived from religions or views of life. According to the first approach, religious beliefs provide believers with reasons for accepting certain moral precepts (principles) *and* with a powerful motivation for living according to those precepts which can, however, be expressed independently of one's worldview and also be endorsed by non-believers or adherents of completely different worldviews. According to the second approach, there is an internal relation between one's moral commitments and the (metaphysical) beliefs, which are part of one's (religious) worldview. It would then be logically incoherent to commit oneself to the way of life expressed in a religious model, i.e. a specific morality, while refusing to believe that what is expressed in this model is somehow factually true. Views of life not merely have a motivational relation to our moral commitments but also determine the content of these commitments and the nature of the actions to which we are committed. This would, for instance, mean that adherents of different views of life or religions are only superficially committed to the same moral actions. Moral principles may prescribe a certain form of observable behaviour, but – as we come to intentions and internal understanding – what is done may be different according to different religious models. This does not make the talk of moral principles inert, however, especially in situations where people have to co-operate in spite of their different worldviews. Though religious beliefs will constitute an essential component within the 'background theories' used to justify moral commitments, reflective equilibrium thinking may take a more or less pragmatic direction in which one aims at reaching moral consensus in terms of shared moral principles and not in terms of the ideological models and metaphors through which people understand the meaning of their lives and experiences.

Carl Reinhold Bråkenhielm further elaborates the relation between what he calls 'cosmological claims', i.e. general beliefs about the ultimate framework of human existence, and the moral and non-moral considerations which play a role in wide reflective equilibrium. He goes into a point that is also touched upon by Ton van den Beld and Vincent Brümmer: cosmological claims cannot be immune to adjustment and modification in the course of the reflective equilibrium process. Though cosmological claims will be clearly underdetermined by other beliefs endorsed in reflective equilibrium (if only for their generality), they are certainly not immune to critical scrutiny. As far as cosmological views claim to make sense of the other beliefs entertained in the equilibrium, their moral and factual implications may be criticised and the cosmological model they stem from can be the object of rational inquiry. The problem remains, however, that even the best methods we have for an assessment of cosmological claims do not give us determinate answers about their truth, so that an element of relativism and perhaps unfounded commitment remains.

This insufficiency of rational argument is an element of any reflective equilibrium, then. It also presents us with the problem of how to give a rational account of major changes in worldviews and major values if these views are seriously underdetermined by argument. In his contribution, Anders Jeffner distinguishes a general social constructivist, a materialistic constructivist, an idealist and an objectivist explanation of changing values and worldviews. He argues that the general social constructivist and the objectivist explanation each have a part of the correct story, which can only be understood if one sees that values are different from valuations. Valuations change, that is for sure, but of values it would be better to say that they are differently and perhaps better understood given changes in valuations. Jeffner argues that a coherentist model of moral thinking and of changes in moral beliefs is very well compatible with a kind of metaethical objectivism.

Part IV, *The Purposes of Reflective Equilibrium: Modelling Public Debate and Dialogue*, discusses various uses of the reflective equilibrium method: as a tool for modelling public debate, an instrument for developing public morality or a way of structuring a (Socratic) dialogue on core values in business. Jan Vorstenbosch discusses the possibility of using the reflective equilibrium method for analysing and organising public debates, i.e. the process of exchanging – usually via the media – opinions, arguments and general ideas about a specific issue whereby this process is accessible, at least in theory, to every citizen. One of the problems of using the method in this way is, of course, that reflective equilibrium is usually thought of as a model of moral thinking by an individual, whereas public debates are essentially collective. Given a plurality of opinions, *whose* considered moral judgements are such that they have to be taken as the starting point of the reflective equilibrium balancing process? Moreover, in public debate not only discussion on merits but also special interests and partiality are the driving forces. What about the conditions for impartiality which are usually considered to be constitutive of moral thinking? Vorstenbosch elegantly shows that the display of a wide range of opinions and arguments in public debates may countervail one-sidedness and overt partiality, and that one could think of a kind of division of tasks for the participants in the debate: intuitions are expressed by lay people, more precisely stated by opinion leaders, corrected by scientists, connected to moral principles formulated by ethicists and brought into interplay with background theories by those who present general views in the course of the debate. According to Vorstenbosch, the method of wide reflective equilibrium has mainly a heuristic function in bringing to light the variety of views on a particular issue (pointing out considerations that otherwise might have been left out of the debate) and making clear how

arguments and considerations of the various positions are related, thereby creating balance and order without excluding viewpoints.

Public debates may play a role in the development of common public morality, i.e. shared convictions on how to interact with the non-specific members of the communities we live in. Frans Brom argues that the wide reflective equilibrium method can be used to understand the emergence and change of this 'morality of the public sphere'. In a pluralistic society, a common public morality is the result of an overlap between the different moralities for the public sphere that are endorsed by different (groups of) citizens. This overlap should not be conceived as a static set of beliefs but as the result of a continuing process of discussion within and between the various moralities, as an overlapping discourse. This overlapping discourse can be characterised as a deliberative interaction between positive social morality (de facto agreement) and critical morality (proposals for amendments and refinement of this social morality). This deliberative interaction may very well be modelled along the lines of an intersubjective reflective equilibrium process.

In the above contributions, reflective equilibrium is presented as a method for pointing out and analysing the conceptual elements of moral thinking (private or public) and the way these elements are related. Jos Kessels argues that we also need to know more about the *methodology* as such, that is, about actual steps, interventions and procedures needed to bring about a reflective equilibrium of considerations. Kessels shows how a modern version of the classical Socratic dialogue can be applied as a tool for achieving reflective equilibrium. He does this by describing an example in which the Socratic dialogue method is used in a concrete case (value conflicts in a business firm). The example illuminates various methodological insights, like the importance of *elenchus*, moments of refutations and embarrassment, and *aporia*, moments of perplexity and deliberative despair, and it shows the fruitfulness of limiting fundamental questions to single examples and narrowing down the thinking process to specific judgements before developing statements of increasing wideness and generality.

In the final part, *Case Studies in Bioethics*, four examples are given of how the reflective equilibrium method can be deployed in so-called 'applied' ethics. Egbert Schroten analyses the case of the bull 'Herman', in which a gene construct, coding for lactoferrine (a protein with possible anti-infectious effects), was successfully introduced. The question is whether it would be morally acceptable to have female offspring from this genetically modified animal in order to obtain lactoferrine in the mammary glands of cows. Schroten was the chairman of the Dutch governmental commission which had to advise on the moral admissibility of the genetic project, and he describes how the various

elements and dynamics of wide reflective equilibrium thinking are visible in the deliberation and advisory work of this (multidisciplinary) commission.

Bart Rutgers applies the reflective equilibrium method to two concrete cases in veterinary ethics: the declawing of cats and the routine practice of Caesarean section in beef cattle. He shows the importance and meaning of three fundamental moral principles that should guide our relationship with animals: non-maleficence, beneficence and the principle of respect for animal integrity. Decision making according to the reflective equilibrium method bears a strong resemblance to the way veterinarians try to solve clinical problems. Therefore, it should not be difficult for a veterinarian experienced in clinical thinking to make himself familiar with this way of moral thinking.

Göran Collste discusses the dramatic issue of infanticide in the case of a mentally retarded, physically suffering newborn infant. He reconstructs various lines of argumentation both in favour of and against the view that the infant should be killed, and focuses on the role of background theories concerning the nature and value of human beings. He argues that also on the level of these fundamental theories and views rational argumentation is possible – e.g. on the importance of human dignity and the value of human life and on the moral importance of suffering – though, finally, the assessment of the arguments may be related to some personal point of view.

Hans van Delden and Ghislaine van Thiel present a normative-empirical model of bioethics which acknowledges the relevance of practice-internal norms for moral reasoning. They show that reflective equilibrium can be used as a method to elicit these internal norms and to integrate the considered moral judgements of professionals with insights and approaches stemming from ethical theory. Their case is the way the principle of respect for autonomy functions in Dutch nursing home care, and how this principle could be interpreted in order to be relevant in the particular context of care for the elderly in nursing homes. While starting from empirical data on the moral opinions of professionals, their normative goal is to suggest a better interpretation of respect for autonomy in nursing-home care. This project suggests fruitful new perspectives for the practical use of reflective equilibrium models and raises complex methodological questions. It is thus a good illustration of a common theme running through many contributions to this book: the fruitfulness of an integrated approach to methodological and pragmatic issues. By trying to deal with practical issues, we learn more about the problems and possibilities of reflective equilibrium methods and, conversely, by using the method of reflective equilibrium we can learn to cope successfully with practical problems.

Ideals and Ideal Theory: The Problem of Methodological Conservatism

Wibren van der Burg¹

1 Introduction

A well-known problem of reflective equilibrium methods is how to avoid or at least correct for a methodological conservatism or, more positively formulated, how to guarantee adequate critical input. If in the process we use only our own convictions, of whatever kind, how can we hope to do more than systematising our prejudices? In other philosophical theories, the common solution is to suggest that there is some Archimedean point or locus of certainty, but this solution is clearly not available for reflective equilibrium methods.

Most strategies to solve this problem in reflective equilibrium methods are based on some form of abstraction or idealisation. We can create, e.g., a distance from our concrete selves, from our concrete society or from our concrete normative judgements regarding the here and now. The strategy of abstraction from our concrete selves and our concrete society is most prominent in Rawls's theory, in the construction of the persons in the original position, who have to make decisions behind a veil of ignorance.

There is yet another way in *A Theory of Justice* in which distance from our own society is created which, in my view, is too often uncritically taken for granted. The normative theory that Rawls develops is not meant to be directly applicable to our own society; it is an ideal theory, that is, a theory for a perfectly just society. A process of transformation is still needed to construct a non-ideal theory – a normative theory for the real society we live in. Rawls, however, does not specify how this transformation process is to take place. He only illustrates it with a very limited non-ideal theory about civil disobedience in a nearly just society. Therefore, the general question remains open under what conditions and how precisely an appeal to ideal theory may or may not be legitimate when constructing a normative theory for the real society we live in or when trying to solve concrete moral problems. My thesis in this article is that the transformation process from a normative theory for an (almost) ideal society to a normative theory for a real society involves so many problems that such an appeal only rarely is valid.

¹ I would like to thank Marcel Verweij en Anton Veldter for their helpful criticisms of an earlier draft.

Therefore, I suggest a third way of creating distance which has close affinities with the appeal to ideal theory, but is less vulnerable to the objections against it. It is the appeal to ideals as such, to valuable future states of affairs. Ideals, like ideal theory, go beyond the principles and considered judgements about what we should do here and now. My argument is that ideals should play a crucial role in reflective equilibrium methods, and that they are especially important as an element which continuously reopens the argumentation processes and the provisional reflective equilibrium to critical perspectives.

It can be defended that ideals play an important role in the construction of the original position as well. Various authors, e.g. Ronald Dworkin (1978) and T. K. Seung (1993), have argued that the original position is based on a normative deep theory in which the (Kantian) ideals of equality and liberty are central. I will not pursue this line – although I think it a valid one – because I want to focus on ideals as substantive elements in the reflective equilibrium process rather than as the basis for methodological requirements of impartiality.

2 The appeal to ideal theory in normative argument

Both in *A Theory of Justice* and in *Political Liberalism*, Rawls stresses that his primary focus is on developing an ideal theory, a normative theory for a perfectly just society. He argues that we should always start with constructing an ideal theory, which subsequently can be the basis for constructing a non-ideal theory.

The concept of ideal theory needs clarification. One possibility, which I will call *descriptive ideal theory*, is to regard an ideal theory as the elaboration of an ideal. For example, it could be a detailed description of the good society we aspire to, or a description of the ideal doctor. I think this kind of theorising about ideals is most useful in normative argument; I will address its use later. Rawls's two principles for the basic structure of society are part of descriptive ideal theory. However, he also goes beyond this. The second part of his ideal theory might be called *normative ideal theory*. This is a normative theory not about the basic normative principles on which an ideal society should be built but on the more specific principles that should guide political action in such a society. Description of a future (or imagined) state of affairs shades into prescription about how institutions and people should act in that ideal context.

If we want to go from descriptive ideal theory to normative theory for the reality we live in, to moral principles that can serve as action guides, we encounter various problems. There is a problem of transformation (how to translate ideals into principles) and one of indeterminacy (how to balance conflicting principles). It seems, nevertheless, possible to construe normative principles. The ideal of peace, for example, can be elaborated into a descriptive

ideal theory about a peaceful world in which no nuclear arms exist and no military threats are made between states. This can then be transformed into normative principles along two lines of argument:² one is teleological and results in a general principle 'act according to the most effective strategy towards this peaceful world'; the other is deontological and results in a general principle 'act according to the principles that would be accepted in this ideal, peaceful world'. The first principle can be elaborated into more precise principles about how to reach the ideal of peace, principles which must be based on an analysis of both the real world and the ideal world. For an elaboration of the second principle, however, we only need to analyse and to elaborate the ideal of the peaceful society to know which specific principles would be accepted in that context. In my view, we need both types of principles in a mixed theory. The crucial question is, of course, how to balance the various principles in the case of conflict, but this is a common problem in moral philosophy. The most important point is that, by accepting the ideal of a peaceful society as a point of reference, we can construct a framework in which all relevant aspects can be fitted, including the realistic assessments of world politics.

If, however, we start from normative ideal theory in order to develop a normative theory for the real world, we encounter more serious problems. The starting point then is a set of specific principles about how political institutions should be structured and how institutions and individuals should act in a future ideal society. Analogous to the second line of argument I distinguished above, these ideal-theory principles can be transformed into deontological principles for our society. However, there is no analogous possibility to derive teleological principles that can adequately deal with the fact that our society is not so ideal and that can lead us from our real life at least in the direction of the ideal. We do not have a mixed normative theory then, but only a theory consisting of deontological principles that have been developed for an ideal situation. Starting from normative ideal theory thus results in a biased and partial normative theory for the real world.

The ideal of peace can serve as an illustration again. Our world is still far from peaceful. The principle that we should never threaten other countries with military force is certainly a defensible principle for an ideal society. It is clear, however, that this need not always be the best principle in the world we live in, for example in dealing with very aggressive neighbours or with countries that commit genocide and other crimes against humanity. Threats or even real military force may be necessary and justified against such countries.

² In section 3, I will show in more detail how we might use ideals to support various types of principles.

So the central problem is clear. Principles that are perfectly justifiable in an ideal society need not be equally justifiable in the non-ideal society we live in. We seem to need at least some teleological principles that show us how to get from our society to the ideal society. Starting from descriptive ideal theory is therefore more fruitful than starting from normative ideal theory.

Does this mean that an appeal to normative ideal theory is never useful? This would be too hasty a conclusion. Normative ideal theory can be the basis for convincing arguments in a special way, which is nicely illustrated by Rawls's own example of normative non-ideal theory: the theory of civil disobedience (Rawls 1971, 333-391).

Starting from his ideal theory, Rawls argues that even in a nearly just society, civil disobedience would be justifiable as a form of protest against certain forms of injustice. Even in such an almost ideal society, civil disobedience should therefore sometimes be tolerated as a legitimate form of democratic protest. If this is a valid argument (which I think it is), we can easily infer that, a fortiori, it is sometimes justifiable and should be tolerated in our own, far from nearly just society. Thus ideal theory offers a strong argument against straightforward law-and-order positions that hold that civil disobedience is never morally justified or politically legitimate.

Rawls's theory of civil disobedience is very restrictive and allows only a limited range of protest actions. In an almost perfect society, such a restriction would certainly be justified. However, this is different in our far from perfect society. The appeal to ideal theory is, therefore, much less convincing against those who argue in favour of a much broader scope for justified civil disobedience. Most acts of civil disobedience are directed precisely against the fact that our current society is not an ideal society. Why then should we use a very restrictive theory, valid for a nearly just society, to judge acts that aim to make our unjust society more just? If our society is much less just, the harmful effects of civil disobedience on social order would probably put less weight into the scales. More importantly, the reasons why we have a prima-facie political obligation in an ideal society – e.g. that the law is largely justified – need not necessarily hold (cf. Van der Burg 1989).

The interesting conclusion is that sometimes an appeal to normative ideal theory can be valid against the arguments of specific opponents, whereas it may not be convincing against the arguments of other opponents in the same debate. In the case of civil disobedience, the appeal is successful against conservative law-and-order positions, but not against more liberal or radical positions that claim a broader justification for civil disobedience. In other cases, the positions may be reversed, and an appeal to ideal theory may be successful against more progressive positions. An example may be income distribution. Rawls's difference principle may be effectively used against radical egalitarians who

claim that everyone should, without exceptions on grounds of merit, earn the same income. It does not help us much, however, once we have accepted that there can be legitimate income differences, in determining how large these differences should be. The difference principle is not effective against the position that there should be no limit to income differences at all because, given the motivations real people have, we must allow them at least the perspective of becoming extremely rich. Here normative ideal theory is of no help.

The conclusion so far is that appeals to normative ideal theory in the context of real-life problems are usually of not much use. It is not clear why normative principles that are justified in the context of an ideal society and presuppose this ideal society for their justification, should be an acceptable starting point for thinking about the normative principles for the here and now. We even have reason to believe that sometimes this starting point will lead to a biased approach.

3 The role of ideals in a network model

The negative conclusion in the previous section does not imply that appeals to ideals and to descriptive ideal theory are equally problematic. On the contrary, the example of the ideal of a peaceful world society suggests that there are various ways in which such an ideal can be legitimately translated into more specific principles for real life.

Let us therefore make a fresh start. I assume that the goal of a reflective equilibrium process is to develop normative theories for the real society we live in, or to reach justified conclusions concerning concrete normative problems. Robert Heeger and Theo van Willigenburg have developed, in a number of publications, a network model of reflective equilibrium (Heeger 1988 and 1990a; Van Willigenburg and Heeger 1991). According to Van Willigenburg and Heeger (1991), a network model consists of three elements: moral principles, moral intuitions and morally relevant facts. In order to reach justified judgements on concrete moral problems, we should try to establish a reflective equilibrium between these three (groups of) elements. The advantages of their model is that it gives due attention to the morally relevant facts and thus to a realistic assessment of the situation and the larger context we have to deal with.

However, in this network model the problem of guaranteeing critical input is particularly urgent. All the normative elements used are directly connected to our own experience. According to Van Willigenburg and Heeger, principles are no more than systematisations of our moral experience, and our moral intuitions are also a reflection of our experience. So how can we go beyond our experience in these models and guarantee critical output?

My suggestion is that introducing a fourth element in the network model, ideals, will be a good way of promoting critical input into the reflective equilibrium process (cf. Van der Burg 1991). Ideals may be defined as follows:³

Ideals are values that are implicit or latent in the law or the public and moral culture of a society or group that usually cannot be fully realised and that partly transcend contingent, historical formulations and implementations in terms of rules and principles.

This definition combines three elements: First, ideals are values rather than direct action guides like principles or rules. Second, an ideal is both grounded in reality and oriented towards the future. Third, ideals are vague and cannot be completely grasped in a formulation, nor can they be completely realised. Especially the third element is a reason why ideals can fulfil such an important role in warranting critical input.

Ideals are much more ambiguous and indeterminate than principles, and can never be fully grasped. Each formulation of a more elaborate conception of an ideal will always describe the ideal only partially. In the light of new situations, we may suddenly see new dimensions of the ideal and reformulate our conception of the ideal. This again may lead to a reformulation of our principles and considered judgements. The fact that ideals transcend every historical formulation thus makes them a continuous source for new critical input into the reflective equilibrium process.

How then do ideals play a role in a network model? The fact that ideals are not direct action guides implies that we need a process of interpretation and transformation in order to connect them to principles or considered judgements. Ideals can be connected to principles along two ways, corresponding to the distinction between deontological and teleological principles.⁴ A first step along both ways should be that from concept to conception,⁵ e.g. from the abstract formulation of an ideal of non-violence to the more specific description that

- 3 I have elaborated this definition in Van der Burg (1997). Taekema (1998) offers an interesting analysis of three ideal-typical models of ideals: cultural, historical and transcendental. My definition is a combination of the cultural and transcendental models.
- 4 Although it is impossible to avoid terms like derivation in this presentation, it should be conceived as a two-way process. We can take the reverse steps to construct ideals on the basis of our principles and intuitions.
- 5 The distinction between concept and conception is central to Dworkin's theory of interpretation (Dworkin 1978) but can also be found in Rawls (1971, 3 ff.). In Rawls (1993, 14), he traces the distinction back to Hart (1961, 155-159).

no one should ever use any physical force against persons.⁶ The second step can either be direct or indirect. A direct derivation is based on the idea that, as much as possible, we should act according to the ideal here and now. The result is a deontological principle, e.g. the principle: 'Do not kill'. An indirect derivation is based on the idea that we should try to realise the ideal in the future, as much as possible, and take those steps that will lead us to that future goal. The result is a teleological principle, e.g. the principle: 'Work towards a non-violent society'.⁷

We can now see how ideals can help to counter the built-in methodological conservatism of Van Willigenburg and Heeger's network version of reflective equilibrium. Because ideals are future oriented and transcend every attempt to formulate specific conceptions, they can help us to go beyond our limited moral experience. They constitute a perspective from which we can look at our own society and our moral views with some distance. They are a continuous source of new normative ideas that can be used to criticise our current principles and intuitions. They are thus, usually in a dialectic interaction with new social or technological developments in reality, one of the major factors in the dynamics of our moral views.

Ideals are certainly no guarantee for adequate critical input; they are just one element in the process which offers no guarantees for success. The recognition of ideals has to be supplemented by other attempts to counter methodological conservatism. One way is to find some guarantees in the person who makes the judgements, e.g. by focusing on the competent judge (as in Rawls 1951) or, following the suggestions in DePaul (1993), by deliberately seeking to enrich our moral experience and thus developing our moral capacities. Another way is to take the words 'our moral experience' seriously in their intersubjective sense and try to introduce intersubjective criticism and discussion. It is not merely my personal experience in which ideals are grounded and which is summarised in moral principles, but the much richer experience of my society, of my culture.

We should nevertheless acknowledge that we will never be fully able to counter methodological conservatism and to take a completely distanced, impartial critical view on our own moral convictions. Despite our attempts, we will always be vulnerable to the criticism that we have not completely gone beyond our own personal cultural biases. But is this really a serious problem

- 6 One could elaborate this conception still further. A more elaborate conception of an ideal would amount to what I have called earlier a descriptive ideal theory.
- 7 Along similar lines, ideals can also directly and indirectly be connected to our moral intuitions. I have more extensively discussed the connection of ideals, principles and policies, rules and concrete judgements in Van der Burg (forthcoming 1999).

or rather an advantage of reflective equilibrium methods? If our aim is to gain moral *knowledge* (whatever this might be), it could be a problem, although we could argue against critics that similar epistemological problems are also unavoidable in the sciences. If our aim is, however, to live a morally good life and to act morally responsibly, we should rather regard the connection to our cultural and personal identities as an advantage. It is this concrete historic person who has to act, not some abstract person in an original position or an archangel. If we want our normative theories or our moral judgements to be effective action guides, they should not be too distanced. A balance should be found between practical effectiveness (people actually living up to their moral views) and moral legitimacy (their moral views being justified). I think a reflective equilibrium which recognises ideals is the best way of striking a balance here: It starts with people's own convictions (including their ideals), and tries to build in as much critical reflection as possible. I do not see how we can do better, being the persons we are rather than some kind of idealised saints.

4 Immanence or transcendence?

So far, I have ignored one major problem of my analysis of ideals: If they are immanent in our culture or experience, how can they also transcend this culture or experience and form a point of critical reflection?

An interesting observation to start with is that many authors who recognise ideals (e.g. Dworkin 1986; Selznick 1961) somehow combine the idea that ideals are immanent in reality or experience, with a claim that they transcend reality. I do not think this is accidental – it is in the nature of ideals that they combine both. The question is not whether these two should be combined, but how.

Seung (1993) has argued that only if we assume a transcendent existence of our fundamental ideals as Platonic forms, can we find a guarantee for full universality of our normative claims. If ideals have no such existence, we can never reach beyond our cultural limitations and attain full universality. The problem with his argument is that, in my opinion, he does not show what such a transcendent existence would be or how we could know it. Nevertheless, his analysis demonstrates an important point: The claim to transcendence is closely linked with the aspiration to universality.

This point should be taken seriously, but can be combined with a constructivist interpretation, which to me seems to be more in line with authors like Selznick, Dworkin and Rawls. The starting point should be that ideals are grounded in reality, in our moral experience. They are latent or immanent values embodied in our culture. They have a basis in experience. For example, they may find their origin in so-called negative contrast experiences, experiences

that something is flagrantly unjust or wrong; the ideal can then arise as an idealised negation of this experience.⁸ The formulation of an ideal of justice may thus be triggered by concrete experiences of clear injustice. Furthermore, the fact that ideals are grounded in experience implies more than merely that they arise from such concrete negative experiences. They may also be supported by positive experiences, for example, when we repeatedly experience that striving for certain ideals is worthwhile, has positive effects and gives a richer meaning to our lives (cf. Rescher 1987).

However, ideals may be grounded in experience, but they cannot be reduced to experience. Their nature is such that, as soon as we recognise them or construct them explicitly, they move us beyond our experiences. Or, perhaps we should formulate this more carefully, in order to avoid the unnecessary suggestion that ideals are independent actors: We move beyond our experiences and beyond our reality as soon as we reflect on the meaning of our ideals. If we formulate a principle of justice, this may be nothing more than the systematisation of our moral experiences; but if we formulate an ideal of justice, we reach a more abstract and open level of reflection. At this level, there is always an indefinite surplus of meaning and an openness to yet unforeseen implications. It is not possible to restrict the meaning of the ideal to the original experiences or principles it systematises; there is always a movement towards broader perspectives. Thus, ideals do not have a transcendent existence but they embody a transcending movement towards more general, more universal meanings.

The conclusion can be that ideals are both immanent in our culture and embody a transcending tendency. I think this reading of ideals in which transcendence and immanence are combined in a coherent way, is the most defensible one – we do not need to suppose some transcendent existence, which would raise many epistemological and metaphysical problems.

I should formulate my point even more precisely, because my interpretation still suggests some form of realism about the existence of ideals in social reality. The suggestion that ideals are latent or immanent in our culture does not mean that they are somewhere out there, merely to be discovered. Our public culture (and even more so the law) is a collective construction, partly grown organically through the ages, partly deliberately created. If I say that a culture embodies certain ideals, this statement presupposes this process of human construction and is itself also the result of my constructive interpretation of that culture. References to the 'immanent existence' of ideals should therefore not be read as reference to a similar form of existence as physical objects have.

8 Cf. Broom (1997, 65); cf. also Van den Brink (1997, 153 ff.).

5 Concluding remarks

In this article, I have argued for the recognition of ideals in reflective equilibrium methods. They can play an important role, especially in the dynamics of reflective equilibrium and in the attempts to get a higher degree of universality and critical reflection on one's own convictions. I have illustrated this thesis with some examples, but I have not given an elaborate demonstration of how this might work in practice; this I have tried to do elsewhere.⁹ The focus in this article was on normative moral or political theory; analogous models of reflective equilibrium can be constructed for legal theory, and in such models, ideals fulfil a similar role (Van der Burg 1993).

The thesis that ideals are especially important for the dynamics of reflective equilibrium and for the enhancement of critical reflection means that their role may not always be the same. If the field of discussion is one in which the moral convictions are relatively settled and stable, the role of ideals will usually be relatively small. The most relevant dimensions of the ideals will then have been acknowledged and formulated in the generally accepted principles. An appeal to ideals to solve concrete problems will then largely be superfluous. If a field is one in which the facts or the moral convictions are in a process of change, e.g. as the result of technological changes, the role of ideals will be much greater. They then provide new perspectives and critical reflection, which may lead to the formulation of new principles and the rethinking of our considered judgements. Similarly, we may expect ideals to be more important in the context of theory construction than in the context of concrete decision making. If we want to develop a normative theory, even a modest one for specific themes such as democracy, ignoring the larger ideals will usually lead to neglecting important dimensions of the theme, if only because many normative phenomena cannot be fully understood unless we acknowledge that they are ideal oriented (Selznick 1961). For problem solving, especially in relatively stable fields, an appeal to ideals may often be much less necessary as long as the central principles connected to these ideals are reasonably justified and generally accepted. Even then, however, the principles will always be open to revision; reflection on new, so far unrecognised dimensions of ideals may easily distort the provisional

equilibrium.¹⁰ It is the process which is central to the method of reflective equilibrium, in my view, not the result. Ideals help to keep the process of critical reflection alive, and hence they help to keep us as humans alive.

9

In Van der Burg (1991), e.g., I applied it to democracy and argued that we cannot understand and justify our considered judgements regarding democracy unless we refer to a larger democratic ideal. At Tilburg University, we have recently started a substantial research programme in which the importance of ideals is investigated, both in theory and in practice, for law, morality and politics (cf. Van der Burg and Brom 1998).

10

The discussion by Marcel Verweij in this volume of how reflection on the ideal of autonomy can lead to the revision of accepted principles is a good example.