

WIDE REFLECTIVE EQUILIBRIUM THEORY  
AND MORAL JUSTIFICATION

by

Ishtiyaque Hussein Haji

B.A., Simon Fraser University, 1981

A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF ARTS  
in the Department  
of  
Philosophy

© Ishtiyaque Hussein Haji 1983

SIMON FRASER UNIVERSITY

March 1977

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

APPROVAL

Name: Ishtiyaque Haji

Degree: Master of Arts

Title of Thesis: Wide Reflective Equilibrium Theory and Moral  
Justification

Examining Committee:

Chairperson: Lawrence Resnick

David ~~Opp~~  
Senior Supervisor

Philip Hanson

-----  
Susan Wendell

Dennis McKerlie  
External Examiner  
Lecturer  
Department of Philosophy  
University of British Columbia

Date Approved: March 31, 1983

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

WIDE REFLECTIVE EQUILIBRIUM

THEORY AND MORAL JUSTIFICATION

Author:

(signature)

ISHTIYAQ HAJI

(name)

Dec. 14, 1983

(date)

ABSTRACT

When is a moral belief or a moral judgement justified? According to Norman Daniels and John Rawls, a moral judgement is justified for a person if it is one that he would accept in "wide reflective equilibrium." Roughly, wide reflective equilibrium is achieved when there is a certain two stage coherence in a person's beliefs: in the first place, between initial moral intuitions or "considered judgements" believed by a person and a set of moral principles; and in the second place, between those moral principles and a set of moral background theories, such as theories of personhood and theories of how one acquires one's morality.

In this thesis, I criticize the appeal to wide reflective equilibrium on three main grounds. First, I argue that neither Daniels nor Rawls seems to have a rationale for using considered judgements as standards against which ethical theory is to be assessed. Second, I argue that, at best, the Rawls/Daniels theory specifies conditions under which persons are justified in holding their moral beliefs, not conditions germane to the justification of the moral principles per se. And third, to the extent that the Rawls/Daniels theory allows different moral principles to be justified for different persons, it may have disturbing implications for the institutionalization of these principles.

I conclude the thesis by proposing a constraint on the choice between competing theories of moral justification. Given two theories of moral justification, other things being equal, it would be reasonable to prefer the theory that provides grounds for justifying a single set of moral principles to be shared by everyone.

ACKNOWLEDGEMENT

I am deeply grateful to David Copp for his persistent encouragement and his invaluable comments and advice. I would also like to thank Phil Hanson, Susan Wendell and David Zimmerman for their comments and suggestions.

WIDE REFLECTIVE EQUILIBRIUM THEORY  
AND MORAL JUSTIFICATION

Table of Contents

Approval . . . . . ii  
Abstract . . . . . iii  
Acknowledgement. . . . . iv

Section

1. Introduction . . . . . 1  
2. Fundamentals: Foundational and Coherence Theories  
Of Epistemic Justification. . . . . 7  
3. The Rawls/Daniels Theory of Wide Reflective Equilibrium. . . 13  
4. Considered Judgements as Assessing Standards . . . . . 30  
5. Objective and Subjective Justification . . . . . 70  
6. Morality and Social Institutions . . . . . 88  
7. Conclusion . . . . . 106  
8. Bibliography . . . . . 108

## Chapter 1

### INTRODUCTION

Two diverse approaches to justifying moral beliefs and theories dominate the current philosophical scene. On the one hand we have the "naturalism" approach, and on the other hand we have the "conservatism" approach.

The naturalism approach identifies moral issues with empirical ones. On this approach, the justification of moral claims is to be settled by methods similar to those used in the justification of other statements about the world. Amongst the various naturalistic theories, Richard Brandt's<sup>1</sup> occupies a prevailing position. In answering questions about the good and the right, Brandt utilizes what he calls the "method of reforming definitions." According to this method, ethical questions are to be settled by first rephrasing these questions in non-moral terminology - in "terminology sufficiently clear and precise for one to answer them by some mode of scientific or observational procedure, or at least by some clearly stateable and familiar mode of reasoning." (1979, p.3) Next, the use of observation and scientific method or some other appropriate form of argument is to establish the acceptability of these "reforming definitions." (1979, p.17) Finally, the use of the appropriate mode of argument is to lead to "a normative conclusion, possibly surprising."

In justifying moral claims, the method of reforming definitions is said to appeal neither to linguistic intuitions, nor to moral convictions, or "intuitions", or "considered judgements" about what is good and right.

In contrast, the conservatism approach typically begins with a consideration of moral intuitions or considered judgements that one has or would have under specified conditions, about concrete situations. In

conservative theories, these considered judgements play an essential role in the development and the justification of a morality, or a theory of moral justification.

Interest in conservatism as an approach to developing and justifying moral theory has been motivated, in part, by its intuitive appeal: It is widespread practice in moral philosophy to establish prima facie credibility for moral principles by showing that the principles are coherent with our moral convictions. In addition many philosophers, like Thomas Nagel,<sup>2</sup> believe that one has no recourse but to start with our moral convictions in justifying moral theory.

In this thesis, my interests lie with the second approach. I will primarily be concerned with the type of conservatism that is involved in the justification of a moral judgement or a morality. David Copp<sup>3</sup> has called this type of conservative theory a type-one conservative theory. He characterizes this type of theory in the following way:

"... a theory of moral justification is (type-one) conservative just in case it implies that a moral judgement's, or a morality's being justified, either simpliciter, or in relation to some person or group, depends on, or is in virtue of, at least in part, that judgement's or that morality's, standing in some appropriate justificatory relation to the considered moral judgements that would be made in specified ideal circumstances, either by everyone, or by that person or that group, and depends on, or is in virtue of, at least in part, their being considered judgements." (p.8)

This type of conservative theory, as developed by John Rawls,<sup>4</sup> and Norman Daniels,<sup>5</sup> has been called into question by, among others, Richard Brandt<sup>6</sup> and David Copp.<sup>7</sup> It has been argued that considered judgements are not credible, and if they are not credible, they cannot be used as standards against which moralities, and theories of the justification of



morality, can be assessed.

I wish to question the tenability of the Rawls/Daniels conservative wide reflective equilibrium theory (hereinafter the 'R/D' theory) in another way. Firstly, I will attack it for its conservatism. Secondly, I will argue that the theory is best construed as a theory specifying conditions under which a person is justified in his moral views, rather than a theory specifying conditions under which a moral judgement or a morality is itself justified. Thirdly, I will question it for its relativism where 'relativism' is to be understood in a particular way.

I will begin, in chapter two, with a brief discussion of foundational and coherence theories of epistemic justification. An understanding of the basic principles of these epistemic theories will put us in a better position to understand and to assess the plausibility of various moral theories of justification.

In chapter three, I will describe the R/D conservative wide reflective equilibrium theory, and some of the problems that may arise with different interpretations of the theory.

In chapter four, I will question the conservatism of the R/D theory on two main grounds: First, I explore why considered judgements might be used as assessing standards in wide reflective equilibrium theories. It is not very controversial that our moral convictions are treated as standards against which morality can be assessed in the wide reflective equilibrium theories of Rawls and Daniels. What is controversial is why these theorists think it appropriate to regard considered judgements as assessing standards. Neither Rawls nor Daniels is explicit on this issue. There appears to be an analogy, if only a rough one, between theory

assessment in science and wide reflective equilibrium theory assessment in ethics. Taking considered judgements to be the rough analogues of observation reports, we might first question why observation reports are used as standards against which scientific theories are assessed, and then consider whether it is plausible that considered judgements can be used as assessing standards for similar reasons. I shall suggest that observation reports can be used as assessing standards if they are:

- (i) taken to constitute the explananda, or
- (ii) they are credible, or
- (iii) they are indicators of reality.

These three criteria are not meant to be mutually exclusive. A credibility story for observation reports, for example, might appeal to realism. Nor do I wish to suggest that these criteria are exhaustive. There may well be additional reasons for using observation reports as assessing standards, but the three criteria I have adduced are the ones that seem to be the most philosophically significant. I will then assess the plausibility of supposing that considered judgements can be used as assessing standards by virtue of being explananda, credible, or indicators of reality. I will argue that considered judgements, unlike observation reports, cannot be used as assessing standards by virtue of being explananda. In addition, I will argue that if considered judgements are used as assessing standards in virtue of being allegedly credible, or indicators of moral reality, then the R/D theory provides a derivative test for the justification of a morality. It provides a derivative test since the very standards against which a morality is to be assessed, considered judgements,

are in need of prior justification.

Second, I will attack the conservatism of the R/D theory by suggesting a number of dilemmas for wide reflective equilibrium theories.

In chapter five, I will distinguish between the notion of a morality's being justified itself, and that of a person's being justified in holding a morality, and I will explain why the former notion is the pertinent one in developing theories of moral justification. I will then argue that the R/D theory is best construed as a theory describing conditions under which a person is justified in holding a moral code, and consequently, it is not a theory germane to the justification of a morality.

Finally, in chapter six, I will contend that the R/D theory implicates the possibility of "intrasocietal relativism". Roughly, this amounts to the claim that the R/D theory allows for the possibility of two or more incompatible moralities being justified relative to different persons within a particular society. I will argue that any theory that implicates the possibility of intrasocietal relativism is liable to encounter difficulties when it comes to the institutionalization of moral principles. I will conclude chapter six by suggesting a constraint that would enable us to choose between competing theories of moral justification. According to this constraint, given two theories of moral justification, all other things being equal, it would be reasonable to prefer the theory that does not implicate the possibility of intrasocietal relativism.

REFERENCES

1. Brandt, Richard B., A Theory of The Good And The Right, Oxford: Clarendon Press, 1979.
2. Nagel, Thomas, "The Objectivity of Ethics", unpublished, 1980.
3. Copp, David, "Considered Judgements And Moral Justification: Conservatism in Moral Theory", forthcoming in David Copp and David Zimmerman, Morality, Reason and Truth, Dordrecht, Holland: Reidel.
4. Rawls, John, A Theory of Justice, Cambridge: Harvard University Press, 1971.
5. Daniels, Norman, "Wide Reflective Equilibrium and Theory Acceptance In Ethics", The Journal Of Philosophy, 76, 1979, pp. 256-282.
6. Brandt, Richard B., *Ibid*, 1979.
7. Copp, David, *Ibid*, 1982.

## Chapter 2

### FUNDAMENTALS: Foundational and Coherence Theories of Epistemic Justification

It will be helpful to begin with a brief consideration of the two main genera of epistemic theories of justification, namely, foundational and coherence theories. This will be useful for a variety of reasons. First, a number of the wide reflective equilibrium models or theories that we will discuss in the next chapter, are closely analogous to foundational epistemic, and coherence epistemic theories. An understanding of the fundamentals of foundational and coherence theories of epistemic justification will put us in a better position to understand theories of moral justification. Second, although there are similarities between foundational and coherence theories of epistemic justification and those of moral justification, there are also significant disparities or disanalogies. A prefatory chapter delineating, amongst other things, the principal characteristics of epistemic coherence theories of justification will facilitate comparison of these theories with coherence theories of moral justification, and hence, will provide a basis for determining whether the alleged disanalogies undermine coherence theories of moral justification.

#### 1. Foundational Theories

Much epistemic justification is inferential in character. The inferential character of epistemic justification gives rise to what is known as the "regress problem".<sup>1</sup> If one proposition is shown to be justified by inferring it from another proposition or set of propositions which are offered as reasons for accepting the initial proposition, the

justifying propositions must themselves be justified. If these justifying propositions are to be inferentially justified, their justification will depend on the logically prior justification of their justifying premisses. If all justification is to be inferential, then this process of justification is threatened by a vicious infinite regress with the apparent result that no proposition is ever justified.

Foundationalism provides one solution to the regress problem. According to the foundationalist, most empirical propositions are justified, if at all, by being inferable from a set of basic propositions, which are "immediately" or non-inferentially justified. The regress of empirical justification terminates when the set of basic propositions is reached, and this set provides a foundation on which all other justification rests.

Different foundational theories arise from different accounts of epistemologically basic propositions and different accounts of the manner in which epistemologically basic propositions epistemically support other propositions.<sup>2</sup>

With respect to epistemologically basic propositions, theories differ in their accounts of which propositions are basic, and in their accounts of the epistemic status of the basic propositions. "Strong" foundational theories, for example, require the basic propositions to be incorrigible, whereas "weak" or "minimal" foundational theories allow the basic propositions to be defeasible.

As far as epistemic support is concerned,<sup>3</sup> two general types of supportive structures or relations further serve to distinguish different foundational theories. A linear structure of support is generally associated with classical foundational theories. On this view, roughly,

proposition P is justified by virtue of some relation which holds between P and another proposition Q, where Q provides justifying reasons for accepting P. Alternatively, the structure of support may be holistic. This type of support is generally characteristic of coherence theories, but foundationalists can also construe support in a more or less holistic fashion. When the supportive structure is holistic, roughly, proposition P is justified by virtue of some relation which holds between P and a whole set of other propositions S, where the propositions in S, and the relations that obtain between these propositions, provide justifying reasons for accepting P.

## 2. Coherence Theories

The essential feature of a coherence theory is that according to this theory, every proposition ab initio, has the same epistemic status as every other. This is the respect in which coherence theories differ from foundational theories, for, according to foundational theories, epistemologically basic propositions have epistemic priority over other propositions.

All epistemic coherence theories share the following features.<sup>4</sup>

- (i) According to these theories, all epistemic justification for empirical propositions is inferential.
- (ii) The regress of justification does not go on for ever but forms a closed loop or a circle that constitutes a system.
- (iii) The basic unit of epistemic justification is such a system, which is justified in terms of internal coherence, i.e., in terms of how each element "fits" or interlocks" with every

other element of the system. The connecting links between the elements in such a system can be one of inference, of mutual confirmation, or of explanation. Presumably, 'mutual confirmation' and 'explanation' are to be explicated in terms of inferential relationships which hold between various propositions of the cognitive system.

Coherence theorists espouse an holistic conception of inferential justification: beliefs are justified by being inferentially related to other beliefs in the overall context of a coherent system. An holistic conception of inferential justification ultimately provides a basis for the solution of the regress problem: It is important, as Nicholas Rescher<sup>5</sup> and Laurence Bonjour<sup>6</sup> have stressed, to distinguish two levels at which issues of justification can be raised, as far as coherence theories are concerned. First, and of secondary importance, the issue may be one of justifying a particular belief, or a small set of beliefs, when the justification of the overall system is taken for granted. At this level, the justification of a particular belief will be linear. No regress is generated since the justification of the system itself is taken for granted. Second, and fundamentally, the issue may be the global one of justifying the cognitive system itself. At this level, the supportive relation between various beliefs is not linear but mutual or reciprocal. Since there is no epistemic priority amongst the various beliefs at this level, there is no basis for a genuine regress. The charge of vicious circularity of justification is denied by appeal to the fact that the justification of particular beliefs ultimately depends not on other particular beliefs but on the system as a whole. According to Bonjour:



"... the fully explicit justification of a particular belief would involve four distinct steps of argument, as follows:

1. The inferability of that particular belief from other particular beliefs, and further inference relations among particular beliefs.
2. The coherence of the overall system of beliefs.
3. The justification of the overall system of beliefs.
4. The justification of the particular belief in question, by virtue of its membership in the system." (p. 287)

Each of these steps depends on the ones which precede it.

Bonjour also enumerates some "essential points" concerning the concept of coherence: First, he claims that coherence is not to be equated with consistency. Coherence has to do with "systematic connections between the components of a system, not just with their failure to conflict." (p. 288) Second, coherence is a matter of degree. According to the coherence theory, for a system of beliefs to be justified, "it must not be merely coherent to some extent, but more coherent than any currently available alternative." (p. 288) This condition seems to necessitate being able to distinguish one cognitive system from another. And third, according to Bonjour, coherence is connected with the concept of explanation: "... the coherence of a system is enhanced to the extent that observed facts can be explained within it and reduced to the extent that this is not the case." (pp. 288, 289)

The details of particular foundational and coherence theories of epistemic justification need not concern us. The basic tenets of these theories, as discussed above, will suffice for our purposes.

REFERENCES

1. For discussion on the "regress problem" see:  
Pampas, George S. and Swain, Marshall (eds.) Essays on Knowledge and Justification, Ithica: Cornell University Press, 1978.  
See also:  
Bonjour, Laurence, "The Coherence Theory of Empirical Knowledge", Philosophical Studies, 30, 1976, pp.281-312.
2. For discussion on the varieties of epistemological theories see:  
Pollock, John L., "A Plethora of Epistemological Theories" in Pampas, George S. (ed.), Justification and Knowledge, Dordrecht: D. Reidel Publishing Company, 1979, pp.93-114.
3. See: Pollock, John L., *Ibid*, and Bonjour, Laurence, *Ibid*.
4. Bonjour, Laurence, *Ibid.*, 1976, p.283.
5. Rescher, Nicholas, "Foundationalism, Coherentism, and The Idea of Cognitive Systematization", The Journal of Philosophy, 71, 1974, pp.695-710.
6. Bonjour, Laurence, *Ibid.*, 1976.

## Chapter 3

### THE RAWLS/DANIELS THEORY OF WIDE REFLECTIVE EQUILIBRIUM

#### 1. The Rawls/Daniels Theory

The R/D theory of moral justification purports to be a conservative coherence theory. It is also a "relational" theory of moral justification. That is to say, according to this theory, moral judgements of a person or a group that are justified, are not justified simpliciter, but are justified in relation to that person or that group.

In describing the R/D theory, it will be helpful to discuss, first, its Rawlsian history, and second, Daniel's contribution to the theory.

In A Theory of Justice<sup>1</sup> Rawls develops what we can call a "reflective equilibrium theory" of moral justification. Although Rawls specifically gives himself to developing and justifying a theory of justice, the ideas presented in the book can readily be extended to the development and justification of moral theory. Two crucial components are constitutive of Rawls' reflective equilibrium theory: a coherence component and a contractarian component. According to Rawls, the justification of a morality is ultimately a matter of coherence - it is "a matter of the mutual support of many considerations of everything fitting together into one coherent view." (p. 579) What are the components that are to fit together in order to yield a justified moral system? Rawls has it that the development of moral theory begins with a consideration of the considered judgements that one would elicit in response to diverse moral issues. A considered moral judgement is a judgement in which one has sincere confidence, and a judgement which one makes or would make without hesitation "under conditions favourable for deliberation and judgement in general." (p. 48) These conditions must be devoid of

distorting factors such as self-interest, inadequate information, abnormal states of mind, and states of emotional upset. (p. 47)

Principles are then adduced to "match" or "fit" these judgements:

"Now one may think of moral philosophy at first (and I stress the provisional nature of this view) as the attempt to describe our moral capacity; or, in the present case, one may regard a theory of justice as describing our sense of justice. This enterprise is very difficult. For by such a description is not meant simply a list of the judgements on institutions and actions that we are prepared to render, accompanied with supporting reasons when these are offered. Rather, what is required is a formulation of a set of principles which, when conjoined to our beliefs and knowledge of the circumstances, would lead us to make these judgements with their supporting reasons were we to apply these principles conscientiously and intelligently." (p.46)

These principles are initially tested by the common method of counter-example. That is, the implications of these principles for particular cases are explored, and the principles are judged to be provisionally acceptable insofar as they agree with our considered judgements for those cases. An acceptable match between principles and considered judgements yields a "narrow equilibrium".

Rawls adds a third tier of components that converts narrow reflective equilibrium to wide reflective equilibrium: the contract apparatus. A selection of the principles of morality, including those of justice, is to be made by choosers placed under a choice situation that Rawls calls the "Original Position". Choosers seek mutual agreement to live together in a society under a single moral code. The moral code that they are to live under is to be selected by them in the choice situation. Rawls places numerous constraints on the choice situation. For example, the choosers must be impartial, instrumentally rational and self-interested. The derivability of the entire contract apparatus,

together with its constraining principles, from numerous moral background theories, is supposed to render it credible as an appropriate mechanism for selecting principles of morality:

"... the conditions embodied in the description of the original position are ones that we do in fact accept. Or if we do not, then perhaps we can be persuaded to do so by philosophical reflection. Each aspect of the contractual apparatus can be given supporting grounds. Thus what we shall do is to collect together into one conception a number of conditions on principles that we are ready upon due consideration to recognize as reasonable." (p.21)

The background theories include theories of personhood and theories of the "well ordered society."

An important element of justification is supposed to be introduced in the move from narrow to wide reflective equilibrium. Moral principles embodied in a narrow reflective equilibrium are derived from purportedly more basic philosophic and normative principles incorporated in the selection mechanism. Since these basic principles are supposed to be philosophically well grounded, principles, such as those in narrow equilibrium, derived from these basic principles, will themselves be well grounded. In summary, according to Rawls, a moral judgement is justified for a person when it is a judgement he would hold if it were coherent with a suitable set of moral principles and background theories.

Daniels<sup>2</sup> contributes to the Rawlsian program by making explicit what is sometimes merely implicit in this program. Like Rawls, Daniels holds that a moral judgement is justified for a person if it is one that he would accept in wide reflective equilibrium. Wide reflective equilibrium is achieved when there is a coherence in an ordered quintuplet<sup>3</sup> of sets of beliefs held by a person: Initial moral judgements believed

by a person are collected and "filtered" so as to include only those of which he is relatively confident, and only those which have been made under conditions which minimize the risk of error. For example, the judgements have been made when the agent (or agents) is calm and in possession of adequate information to judge the cases. Let the alphanumeric character  $a_1$  represent the "filtered" set of initial moral judgements. The set ( $a_1$ ) of considered moral judgements is to match with a set (p) of moral principles in order to attain a "level I" partial equilibrium. Since various sets of moral principles could have "varying degrees of 'fit' with the moral judgements" in ( $a_1$ ) (p.258), a set ( $b_1$ ) of moral background theories, such as theories of personhood, must show that principles in (p), "level II" principles, are more acceptable than alternatives. Since the theories in ( $b_1$ ) are moral theories, it is likely that they will be constrained by considered judgements. If the same set of considered judgements as those included in the level I partial equilibrium, were to constrain the background moral theories, then inclusion of this component in the coherent system would be superfluous. As Daniels<sup>3</sup> contends, unless the theories in ( $b_1$ ) show that principles in (p) are more acceptable than alternatives on grounds somewhat independent of (p)'s "fit" with relevant considered judgements (p.86), principles in (p) would have no support they would not already have had at the level of partial equilibrium. In addition, unless the theories in ( $b_1$ ) are constrained by a set of considered judgements to some degree independent of the set included in level I partial equilibrium, one might be concerned about "rigging" in the coherence system. That is, Rawls and Daniels might not be able to escape the charge that the background theories are rigged

to yield principles in accord with our pretheoretical moral convictions. For these reasons, Daniels holds that if the moral background theories in  $(b_1)$  are to cohere with a set of considered judgements  $(a_2)$ , then  $(a_1)$  and  $(a_2)$  must be disjoint to a significant degree. Finally, a set  $(b_2)$  of empirical theories, "level IV" theories, such as general social theories, and theories of moral development, are to constrain the set of moral theories in  $(b_1)$ .

Daniels essentially sees Rawls' contract apparatus as "a feature of a particular wide equilibrium." (p.87) The contract apparatus, he claims, is not self-evidently acceptable. "Rather, philosophical argument must persuade us it is a reasonable device for selecting between competing conceptions of justice (or right). These arguments are inferences from a number of (Level III) background theories ..." (p. 87) Daniels concedes that different derivational mechanisms would be acceptable if they were derivable from a set of background moral theories preferable to the theories advanced by Rawls (p. 261).

It is evident that according to this theory, it is not sufficient for a considered moral judgement to be justified that it merely be consistent with a set of moral principles and background theories. The judgement, in order to be justified, must stand in some complex coherence relation R with background theories and principles. In addition Daniels advocates extensive revision of considered judgements. Indeed he holds that the set of considered judgements that will cohere with the rest of the equilibrium package may emerge only when reflective equilibrium is reached, and these judgements may still undergo revision in the light of further theory change. The revision in question might be theory-based

or it might not be so based. To say that revision is theory-based is to say that theoretical considerations undermine (or enhance) the confidence with which an agent believes a considered judgement.

## 2. Different Models Of Wide Reflective Equilibrium

Various interpretations or models of wide reflective equilibrium are possible. In this section, I intend to describe and discuss some of these models. Since I am primarily interested in conservative theories, I will not discuss non-conservative wide reflective equilibrium theories.

We can distinguish different models of conservative wide reflective equilibrium theory on the basis of: (i) whether the theory is a foundational or a non-foundational theory; (ii) whether or not the theory takes considered judgements, under specified conditions, to be indicators of moral facts, and (iii) whether the theory permits extensive or minimum revision of considered judgements. We will discuss each of these categories in turn. Needless to say, a certain theory may exhibit characteristics of more than one of the categories. For example, a foundational conservative theorist might not permit extensive revision of considered judgements that he takes to be indicators of moral reality.

To begin, first, we distinguish between foundational and non-foundational conservative wide reflective equilibrium theories. According to foundational conservative theories, considered judgements, under specified conditions, are "basic" or self-justified. Such theories resemble foundational theories of epistemic justification.

There can be different types of foundational conservative theorists depending on their account of the manner in which basic moral judgements support other moral judgements.



Non-foundational conservative wide reflective equilibrium theorists do not take considered judgements, under specified conditions, to be basic. These theorists resemble coherence epistemic theorists, and like these epistemic theorists, they may put forward different accounts of how each element within the coherence system is supposed to "fit" or "interlock" with every other element of the system. Daniels, for example, holds that philosophical argument is to be adduced to decide between competing sets of principles that cohere with the initial set ( $a_1$ ) of considered moral judgements. These arguments, in turn, according to Daniels, can be construed as inferences from a set of relevant background theories.

The problems that are likely to affect foundational conservative wide reflective equilibrium theories will be analogous to the problems that affect foundational epistemic theories. To name only two, such theorists will have to account for how the basic set of judgements are themselves justified, and they will also have to confront the problem of clearly defining a class of basic judgements.

Non-foundational conservativists, if they are coherence theorists, will encounter the problem of developing a viable coherence system of moral justification. If we take the R/D theory to be a conservative coherence theory, then chapters (4) and (5) are attempts to point out the inadequacies of this theory.

Second, we can distinguish different models of wide reflective equilibrium on the basis of the moral ontology presupposed by the theories. Some conservative theorists might utilize the naturalistic model of wide reflective equilibrium. According to the naturalistic model, considered judgements, under specified hypothetical circumstances,

are indicators of moral reality. Ronald Dworkin characterizes the naturalistic model in the following way. Dworkin<sup>5</sup> asserts that according to this model, considered judgements are

"clues to the nature and existence of more abstract and fundamental moral principles, as physical observations are clues to the existence and nature of fundamental physical laws" (p.160)

This model postulates the existence of moral facts and the method of equilibrium can be viewed as a means of discovering these facts. Since considered judgements held in equilibrium are taken to be indicators of these facts, and, consequently, are likely to be true, they are taken to be justified.

Theorists who deny the existence of moral facts will not utilize such a model. Let us call such theorists non-naturalistic, or non-realist, conservatives.

Advocates of the model of wide reflective equilibrium for theory justification in ethics (for example, Daniels and Thoma Nagel<sup>6</sup>) generally seem to believe that there is only one overall model for the justification of both moral and non-moral beliefs, namely their model. Even if this claim is granted, it is important to realize that there can be different models or interpretations of wide reflective equilibrium, and different theorists both in ethics and in science, may utilize different models. In scientific practice, for example, observation reports can be taken to be rough analogues of considered judgements. In so far as the theorist assumes that observation reports are indicators of non-moral reality, he adopts a naturalistic model of wide reflective equilibrium.

The naturalistic model, both in science and in ethics, enables an

understanding of how observation reports or considered judgements, can assume the role of assessing standards. They are to be treated as assessing standards because they are taken to be indicators of reality; theories in wide reflective equilibrium, according to this model, are to be assessed against the world via considered judgements or observation reports.

It is incumbent upon theorists who deny that considered judgements, under specified conditions, are indicators of reality, to specify why considered judgements are to be treated as assessing standards. Such theorists might contend, for example, that considered judgements are assessing standards because they are credible, leaving it open that an account of the credibility of considered judgements might be given without invoking realism. We will have a lot more to say on the issue of why theorists might be inclined to treat considered judgements as assessing standards in wide reflective equilibrium theories in the next chapter.

Third, we can distinguish between conservative wide reflective equilibrium theorists who permit extensive revision of the initial set of considered judgements and theorists who do not. If these theorists are coherence theorists, "permission" is to be understood in terms of basic methodology: Judgements and theories are to be revised until they "fit" together in a coherent system. Let us call the former type of theorists constructivists or revisionists, and the latter type of theorists non-revisionists. We can also distinguish between type 1 and type 2 revisionists. Type 1 revisionists permit extensive revision of both the initial set of considered judgements and those considered judgements held in wide reflective equilibrium. Perhaps they do this because they wish to distance themselves from foundational theorists. In addition, revision of judgements

in equilibrium may be justified on the basis that equilibrium is "dynamic". Such theorists may claim that revision of judgements in equilibrium may serve to improve overall coherence. On the other hand, type 2 revisionists restrict revision to the initial set of considered judgements, i.e., to the set of judgements held prior to achieving equilibrium. I take Norman Daniels to be a type 1 revisionist theorist.

Having characterized three broad divisions among conservative wide reflective equilibrium theory, namely, foundational or non-foundational, naturalistic or non-naturalistic, and revisionist or non-revisionist, it is interesting to note that some conservative wide reflective equilibrium theories, exhibiting certain combinations of characteristics across divisions seem prima facie to be untenable. I will discuss two such theories, a naturalistic/revisionistic theory and a foundational/revisionistic theory.

Suppose, for example, a conservative theorist adopted the naturalistic model, and he permitted extensive revision of both the initial sets of considered judgements and those considered judgements held in wide reflective equilibrium. Since this theorist is a naturalist, he takes considered judgements, under specified conditions, to be indicators of moral fact. Let us suppose this theorist contended that considered judgements, perhaps under specific conditions, are indicators of fact prior to achieving equilibrium, and he supported this contention by sound independent reasoning. Given that these judgements are indicators of fact, if this theorist were now to permit extensive revision of this initial set of judgements, such revision would be tantamount to rejecting truth. For example, if at time  $t_1$ , one's beliefs or considered judgements

indicated facts  $f_1$ ,  $f_2$ ,  $f_3$ , and at time  $t_2$ , after revision, one's beliefs indicated different facts, then it would appear that at different times, considered judgements would reflect different truths. And furthermore, the example suggests that the facts that considered judgements reflected at  $t_1$ , would be discarded. Of course, the theorist could always claim that at time  $t_1$ , he had made a mistake: He would have to claim that prior to revision, he had mistakenly believed that considered judgements indicated reality. But were he to do this, then, allowing for the possibility of revision after  $t_2$ , it would seem that rather than claiming that considered judgements are indicators of fact, the theorist should more appropriately claim that he believes that considered judgements are indicators of fact. But in this case, theories in wide reflective equilibrium would be assessed not against moral facts, but against psychological states, namely, beliefs of theorists. And, on the face of it, this seems untenable; it is not normally the case that we take theories to be assessed against the mere beliefs of agents.

If, on the other hand, the theorist contended that considered judgements indicated moral facts only in equilibrium, then once equilibrium obtained, he would have little, if any, reason to extensively revise these judgements. Again, revision of these judgements, under these conditions, would be analogous to rejecting truth. It would seem that a conservative theorist would do best were he to claim that it is only considered judgements in wide reflective equilibrium that are indicators of fact, and were he to restrict revision to the initial set of considered judgements. Such a theorist could permit extensive revision of considered judgements in reaching equilibrium, but once equilibrium has been reached, he would permit only minimal revision. If he were to do this, he would in effect

be endorsing the position that it is only considered judgements in wide reflective equilibrium that can be used as assessing standards. However, in chapter (4), when I discuss the "Revisionist's dilemma", I will argue that it is implausible for a conservative theorist to take a set of considered judgements to be legitimate assessing standards if and only if in wide reflective equilibrium.

The conservative might reply that the above objection assumes that a judgement's being an indicator of reality implies that it indicates or reflects truth. This assumption, he might claim, is unwarranted, since an indicator of reality need not be an infallible indicator. The conservative might hold that it suffices that we take it that considered judgements are indicators of moral reality in that one is always reasonable to believe a considered judgement.

The plausibility of this reply depends on how one is to construe "reasonable". For example, with physical theory, it can be argued that one is reasonable to believe that an observation report indicates reality in that causal explanations are available to this effect. But what account of reasonability does the naturalistic conservative offer? In the absence of any plausible account, the conservative's reply loses force.

As a second example, consider a foundational/revisionistic theory. According to such a theory, a set of considered judgements is to be used as an assessing standard in virtue of its members being basic or self-justified. Suppose that the initial set ( $a_1$ ) of considered judgements were considered to be basic. Then, extensive revision of this set would be tantamount to altering, and hence rejecting, the very standards against which theories are to be assessed. If, for example, having established

that judgements  $j_1, j_2, \dots, j_n$  were basic, and if one were to revise these judgements thereby obtaining judgements  $a_1, a_2, \dots, a_n$ , then one would be assessing theories against standards that were not legitimate assessing standards, since according to the theory, it is  $j_1, j_2, \dots, j_n$  that are the basic judgements.

It will not do for the conservative to reply that basic judgements are not necessarily incorrigible and therefore not immune to revision: The most plausible type of revision is theory-based. But if allegedly self-justified judgements can be revised by theoretical considerations, then it would seem that these judgements are not self-justified. This is because their status as allegedly basic judgements would then depend on theoretical considerations.

Suppose, on the other hand, the theory specified that it is only considered judgements in wide reflective equilibrium that are basic, and consequently, that only this set of considered judgements should be used as assessing standards. In this case, the theory would have to restrict revision to the initial set of considered judgements. It appears prima facie that the most plausible version of a foundational/revisionistic theory is a version according to which only those considered judgements in wide reflective equilibrium are to be taken as basic, and according to which revision is to be restricted to the initial set of considered judgements.

### 3. A Possible Inconsistency In The Rawls/Daniels Theory

It is clear that Daniels is a revisionist conservative theorist: he permits extensive revision of both the initial set of considered judgements that one holds and the set of considered judgements that one

would hold in wide reflective equilibrium. On the issue of whether considered judgements are indicators of moral fact, Daniels' position is that:

"It is plausible to think that only the development of acceptable moral theory in wide reflective equilibrium will enable us to determine what kind of 'fact', if any, is involved in a considered moral judgement." (p.271)

In addition, Daniels purports to be a coherence theorist and he finds it unattractive to grant to considered judgements, or to moral principles, a privileged epistemological status (p. 257). This suggests that Daniels wishes to adopt a theory that is non-foundational and type 1 revisionistic.

Before proceeding to the next chapter, we might pause to consider a possible problem that arises for a theorist like Daniels who alleges that he is a coherence theorist and who utilizes a conservative wide reflective equilibrium model for theory justification in ethics.

As contended in the last chapter, an essential feature of a coherence theory is that such a theory initially assigns to every belief or proposition the same epistemic status as it assigns to every other belief. In conservative wide reflective equilibrium theories, considered judgements assume the role of assessing standards. But if such judgements are to assume the role of assessing standards, they must have a different epistemic status from moral judgements that are not considered, i.e. according to such theories, considered judgements will be epistemologically privileged. But if this is the case, then there appears to be an inconsistency in being a coherence theorist who uses considered judgements as assessing standards. In fact, if considered judgements are to be used as assessing standards, and if considered judgements have epistemic



priority over moral judgements that are not considered, then considered judgements in wide reflective equilibrium theories can be regarded as providing a foundation on the basis of which all justification rests. This foundation need not be a "strong" foundation: As we have seen, there is no incompatibility in being a foundational theorist and in permitting extensive revision of the initial set of considered judgements. We suggested above that the most plausible version of a foundation/revisionistic theory is a theory which confines revision to the initial set of considered judgements, and which holds that it is only those considered judgements in wide reflective equilibrium that are to be used as assessing standards, in virtue of being basic. In summary, I am claiming that it is more appropriate to regard Daniels as a weak foundational/revisionistic theorist than as a non-foundational/revisionistic theorist.

The import of this criticism is not merely terminological, i.e., it does not merely suggest a reclassification of Daniels as not being a "pure" coherentist. This criticism, if cogent, bears on two important issues. First, I suggested that a theorist who is a foundational/revisionist would most plausibly contend that it is only considered judgements in wide reflective equilibrium that are to be used as assessing standards. However, in chapter (4) I will argue that any theory committed to this position runs afoul of the "Revisionist's dilemma". And second, the criticism suggests that Daniels' theory of justification is a derivative theory: Daniels must presuppose, prior to applying his criterion of justification, that considered judgements as a type, are suitable candidates for assessing standards; it must be presupposed that considered

judgements possess an epistemic priority over non-considered judgements. A rationale will be required to justify the use of considered judgements as assessing standards. But if the very standards against which a morality is to be assessed are in need of independent justification, then this suggests that conservative wide reflective theories are derivative theories of justification. I suspect they will be derivative in the sense that most of their justificatory force will derive from the independent theorizing required to establish that considered judgements themselves can be adequately treated as assessing standards. The next chapter will assume the burden of at least partly substantiating these claims.

REFERENCES

1. Rawls, John, A Theory of Justice, Cambridge: Harvard University Press, 1971.
2. Daniels, Norman, "Wide Reflective Equilibrium and Theory Acceptance In Ethics", The Journal of Philosophy, 76, 1979, pp.256-282.
3. Copp, David, "Considered Judgements And Moral justification: Conservatism In Moral Theory:", forthcoming in David Copp and David Zimmerman, Morality, Reason and Truth, Dordrecht, Holland: Reidel, 1982.
4. Daniels, Norman, "Reflective Equilibrium and Archimedean Points", Canadian Journal of Philosophy, 10, 1980, pp.83-103.
5. Dworkin, Ronald, Taking Rights Seriously, Cambridge: Harvard University Press, 1977.
6. Nagel, Thomas, "The Objectivity of Ethics", unpublished, 1980.

## Chapter 4

### CONSIDERED JUDGEMENTS AS ASSESSING STANDARDS

It is common practice in moral philosophy to assess specific moral principles by considering whether such principles cohere, or "match", or "fit" with our moral convictions. A "mismatch" between principle and conviction gives us prima facie grounds for impugning the principle in question. That this is so lends credence to the view that considered judgements in conservative theories of ethics assume the role of assessing standards. John Rawls acknowledges this when he affirms that

"There is a definite if limited class of facts against which conjectured principles can be checked, namely, our considered judgements in reflective equilibrium." (p.51)<sup>1</sup>

It is not very controversial that our moral convictions are treated as standards against which moralities can be assessed in the wide reflective equilibrium theories of Rawls and Norman Daniels.<sup>2</sup> What is controversial is why these theorists think it appropriate to regard considered judgements as assessing standards. Neither Rawls nor Daniels is clear on this issue.

In this chapter, I wish to do two things. First, I wish to explore why considered judgements might be used as assessing standards in wide reflective equilibrium theories<sup>1</sup>. I adopt the following strategy: There appears to be an analogy, if only a rough one, between theory assessment in science and wide reflective equilibrium theory assessment in ethics. Taking considered judgements to be the rough analogues of observation reports, we might first question why observation reports are used as standards against which scientific theories are assessed, and then

consider whether it is plausible that considered judgements can be used as assessing standards for similar reasons. I shall suggest that observation reports can be used as assessing standards if they are:

(i) taken to constitute the explananda, or, (ii) they are credible, or, (iii) they are indicators of reality. These three criteria are not meant to be mutually exclusive. A credibility story for observation reports, for example, might appeal to realism. Nor do I wish to suggest that these criteria are exhaustive. There may well be additional reasons for using observation reports as assessing standards, but the three criteria I have adduced seem to be the most philosophically significant. I will then assess the plausibility of supposing that considered judgements can be used as assessing standards by virtue of being explananda, or credible, or indicators of reality, and if they are used in any one of these ways, what the implications are for theory assessment in ethics if one utilizes a wide reflective equilibrium model.

Second, I will question the tenability of the Rawls/Daniels theory of moral justification by suggesting two dilemmas for this theory, the "Conservatives's dilemma" and the "Revisionist's dilemma". The conservative's dilemma affects all conservative theories of wide reflective equilibrium, whereas the Revisionist's dilemma is limited in scope, and affects only revisionist conservative wide reflective equilibrium theories.

#### 1. Considered Judgements As Explananda

Observation reports might be used as "assessing standards" against which scientific theories are assessed if the observation reports are taken to be the explananda. But this sentence is ambiguous. It is open to at least three interpretations: First, we can take the sentence to

imply that the explanandum is a psychological state, i.e., the explanandum is the belief expressed in the report. Second, we might interpret the sentence as claiming that what is reported in the report is the explanandum. And third, the sentence can be interpreted as implying that the explanandum is a speech act - it is the making of the observation report that is the explanandum. Let us consider the plausibility of each of these interpretations.

According to the first interpretation, the explanandum is the belief expressed in the report. This interpretation is dubious since we do not normally take our scientific theories to be assessed against the mere beliefs of people, even intelligent people, but against what there really is in the world. We take our scientific theories to be assessed against reality. This suggests that the second interpretation is more promising.

According to the second interpretation, what is reported in the reports is the explanandum. But if this is true, the reports must be taken to be veridical in order to be explananda and this seems to collapse the distinction between using observation reports as assessing standards in virtue of the reports being explananda, and in virtue of the reports being indicators of reality. Nevertheless, this interpretation seems credible.

We can attest to its credibility by observing the part played by what is reported in observation reports in accounts of scientific explanation. According to a popular, although controversial theory of explanation, the deductive nomological theory, an event or regularity is explained if the proposition that it occurs can be deduced from appropriate covering laws and statements of fact. In turn, regularities that can be

expressed in the form of empirical laws can themselves be explained by subsumption under theoretical laws.

Typically, a good theory will offer a systematized unified account of diverse phenomena. Newton's theory of motion and gravitation, for example, as Hempel indicates,<sup>3</sup> accounts for a number of regularities such as those exhibited by free falling objects, the simple pendulum, the motion of the planets, the tides, etc.

This systematizing or unifying characteristic of a good theory enables us to give another sense to the notion that what is reported in observation reports can be used as assessing standards: In the first way, they can be used as "assessing standards" simply because they can be subsumed under theory; the theory explains the phenomena we are interested in. The terms "assessing standard" when used in the first way can be misleading. Perhaps here we should say that the events or the regularities that observation statements report furnish the data on the basis of which a theory is constructed. In the second way, given the systematizing-unifying function of theories, if a theory  $T_1$  accounts for a larger range of phenomena than an alternate theory  $T_2$ , other things equal,  $T_1$  is to be preferred to  $T_2$ . The explananda force the choice, so to speak, between  $T_1$  and  $T_2$ .

Before we consider the third interpretation of the claim that observation reports are explananda, let us examine the analogous situation in moral theory, where what is reported by considered judgements is the relevant explanandum. That is to say, in moral theory, theorists might claim that considered judgements are treated as assessing standards because what is reported by the judgements is taken to constitute the

explanandum. But in making such a claim, the theorist must presuppose moral realism. He must, for example, presuppose that there is a fact of the matter as to whether capital punishment is wrong. However, in the absence of good arguments for moral realism, a theorist would not be justified in presupposing the existence of moral facts, and consequently, he would not be justified in using considered judgements as assessing standards if his rationale for doing so were that considered judgements are to constitute the explananda. Of course, a wide reflective equilibrium theorist can hypothesize that moral realism is correct, and hypothesize that considered judgements, under specified conditions, are indicators of fact. On the basis of these hypotheses, he could then theorize that considered judgements are to be treated as assessing standards. But in so far as he does this, the legitimate use of what considered judgements report as explananda will be contingent on whether moral realism is true. And unlike non-moral realism, the issue of whether moral realism is true is very controversial.

Regarding the second option as to what constitutes the explananda, a theorist can deny that in taking what is reported by considered judgements as the explananda, he must presuppose moral realism. He can deny this by reminding us about phenomenalism in epistemology. According to the phenomenalist, all talk about material objects can be reduced, without loss of content, into talk about sensations, or *sensa*, or perceptions. The phenomenalist, unlike the physicalist, can contend that theories in science are to be assessed against *sensa*; they are to be assessed against *sensa* since the contents of observation reports are nothing over and above *sensa*.



In an analogous way, the moral theorist may wish to espouse a form of "moral phenomenalism". He may, for example, claim that talk about what is reported by considered judgements can be reduced without loss of content into talk about "moral sensa". But such a theory would be incredible since it would postulate the existence of unobservable, unknown and mysterious moral sensa. An alternate, more plausible theory of "moral phenomenalism" would be a theory according to which all talk involving moral predicates could be reduced, without loss of content, into talk about the attitudes of an ideal observer. For example, such a theory might hold that "X is good" means "If there were an ideal observer, he would have a pro attitude towards X."<sup>4</sup> Roderick Firth,<sup>5</sup> who develops a version of ideal-observer theory, lists the following as essential features of the ideal observer: He is omniscient regarding non-ethical facts, he is omniperceptive, he is disinterested, he is dispassionate, he is consistent and he is normal in other respects.

An advocate of ideal-observer theory may suggest that moral theories in an equilibrium are to be assessed against the attitudes of an ideal observer. The relevant explananda would be the attitudes of an ideal observer.

However, if what is reported by considered judgements are the attitudes of ideal observers, then moral theory need not figure in an explanation of these attitudes. What would explain the attitudes of ideal observers need not be moral theory, but psychological and sociological theory, together with the constraints imposed on what it is to be an ideal observer. It is logically possible that different ideal observers, despite the constraints they must satisfy in order to be ideal observers, could

disagree about moral questions. Different ideal observers could have different attitudes towards the same moral issue. An ideal observer, for example, brought up in a religious environment might have a negative attitude towards abortion, whilst some other observer, brought up under different conditions, might have a positive attitude. Moral theory would be relevant to explaining the attitudes of ideal observers, if the best explanation of their attitudes presupposed the existence of moral facts which causally influenced their attitudes. But it is not open to a "moral phenomenalist" to argue in this way - one cannot both be a moral phenomenalist and espouse moral realism.

So far we have considered two interpretations of the proposal that observation reports are to be used as assessing standards if they are taken to be the explananda. We have argued that considered judgements cannot be used as assessing standards for similar reasons. Now, we turn to the final interpretation. According to the third and final interpretation of the claim that observation reports are to be used as assessing standards in virtue of being explananda, it is the making of the observation report that is the explanandum. This interpretation is also plausible, but its plausibility depends on presupposing non-moral realism. This needs to be explained.

Consider the following: A physicist expounds a certain theory according to which protons exist. Can the making of an observation report be used as confirming evidence for the theory?<sup>6</sup> If it can, then the making of an observation report can be used as a standard against which scientific principles are assessed.

In science, it appears that the making of an observation report can

be used as confirming evidence for theoretical principles. To see this consider how our physicist would go about testing the principle that protons exist. Given his theoretical knowledge, let us suppose that he sets up an appropriate experiment, a cloud chamber experiment. Let us adopt Gilbert Harman's<sup>7</sup> suggestion that we can take an observation in the relevant sense to be "an immediate judgement made in response to the situation without any conscious reasoning having taken place." (p.6) Seeing a vapour trail in the cloud chamber, our physicist makes the following observation: "There goes a proton." The making of the observation is confirming evidence for physical theory which postulates the existence of protons. It is confirming evidence since physical theory explains why the scientist, in the relevant situation, makes the observation that he does make and explains this in the following way: The theory postulates the existence of protons and explains the meaning of the term 'proton'. The proton explains the appearance of the vapour trail; and the vapour trail together with facts about his psychology and physiology explains why the scientist makes the observation that he does make.

In summary, physical principles like the principle that protons exist, are tested against the world via the making of observation reports, under relevant conditions. The making of observation reports can be used as evidential support for or against physical theory since physical facts postulated by physical theory provide the best explanation for the making of these reports. Consequently, the making of observation reports can be used as assessing standards for physical theory.

Consider, now, the moral case. In wide reflective equilibrium theories, it might be claimed that considered judgements are to be

treated as assessing standards because the making of these considered judgements constitutes explananda. In a certain society, for example, it will be discerned that the members of the society display certain moral behaviour. Under specified circumstances, for instance, it might be noted that the members are disposed to make the judgements, "Stealing is wrong", "Abortion is wrong", "Capital punishment is wrong". A theory is sought that explains why the members, under specified conditions, make the judgements that they do make. The specified conditions would include such conditions as the judgements having been made when the agents were calm and in possession of adequate information.

In Daniels' revisionist model of wide reflective equilibrium, a considerable revision of the initial set of judgements is permitted. We are to imagine the theorist working back and forth, making pertinent adjustments to background theory, principles, and considered judgements until an equilibrium point is reached. The fact that Daniels permits extensive revision of the initial set of considered judgements makes it implausible to suppose that the making of this set of considered judgements in the revisionist model is to be construed as explanandum. To suppose this would simply amount to diverting attention to the making of an irrelevant set of considered judgements.

A wide reflective equilibrium theorist, however, may wish to claim that it is only the making of considered judgements in wide reflective equilibrium that is to be used as explanandum. However, a number of problems may arise if the theorist wishes to adhere to this position. First, at the outset of theorizing, a theorist might assume a high degree of "foundationalism", that is to say, he might assume that his morality

is more or less correct. If we are to regard the wide reflective equilibrium program as a program of moral explanation, such a theorist might argue that it is the making of the initial set of considered judgements that require explanation and not the making of some alternate set that emerges in equilibrium. However, it must be noted that nothing compels a theorist to assume a high degree of "foundationalism". But secondly, if a theorist holds that it is only the making of considered judgements in equilibrium that is to constitute the explanandum, then he must hold that it is only the explananda in equilibrium that can be used as assessing standards. But if he does, I will argue, towards the end of the chapter, that he runs afoul of the Revisionist's dilemma.

There are a number of considerations that militate against taking the making of considered judgements as explanandum. The most serious of these considerations, as Harman<sup>8</sup> indicates, pertains to the fact that moral theory (or moral principles) need not figure in an explanation of why people make the considered judgements, or the moral observations, that they make. In other words, unlike the situation in science, moral observations, unless moral realism is true, need not provide evidence for or against moral theory or moral principles. To see this, consider the following example of Harman's:<sup>9</sup> Rounding a corner, I see some children pour gasoline on a cat and ignite it. I spontaneously make the moral observation, O, "The children are wrong to set the cat on fire." What moral principle, if any, explains my making this observation? It might be thought that the following principle, P, suffices: "It is wrong to cause unnecessary suffering."

However, principle P "explains" why it was wrong of the children

to set the cat on fire, but it does not explain my making the observation O which is the relevant explanandum. What is needed to explain my making observation O is my acceptance of principle P or my belief in P and not P itself. My acceptance of P or my belief in P, in turn, can be explained by psychological or sociological mechanisms. Unlike the case in science, where one has to make assumptions about physical facts in order to explain the occurrence of physical observations that support scientific theory, in the moral case, one does not need to assume anything about moral facts to explain the occurrence of moral observations. The occurrence of moral observations can be explained by assumptions about the psychology or moral sensibility of the person making the observation. But if this is true, then moral theory need not be required to explain why a person makes the moral observations that he does make. Consequently, if the making of moral observations or considered judgements are to be used as assessing standards in virtue of being explananda, it is unlikely that we will succeed in justifying a morality.

There is a second major reason why one might be reluctant to construe the making of considered judgements as explanandum in a theory of wide reflective equilibrium. A wide reflective equilibrium theory is a theory of moral justification; it is a theory that purports to delineate conditions under which a moral judgement is justified for a person. If the making of considered judgements in such a theory were to be treated as explanandum, then wide reflective equilibrium theory would be an explanatory theory of moral behaviour. However, there is a distinction between construing a wide reflective equilibrium theory as a theory of the justification of a morality and as a theory of the explanation of moral behaviour. That our

making considered judgements can be explained no more justifies them than that the fact that our astrological judgements, or our judgements about witches can be explained shows them to be justified.

Explanation would be relevant to justification only if the best explanation of our making considered judgements, under specified conditions, presupposed that considered judgements were indicators of reality. If considered judgements were indicators of moral fact, then under specified conditions, one would be able to explain the making of a considered judgement by appeal to the relevant moral fact, just as in the non-moral realm, one can explain why a physicist, under relevant conditions, makes the observations that he does make, by appeal to physical facts.

In summary, it is reasonable to conclude, in the absence of good arguments for moral realism, that considered judgements cannot assume the role of assessing standards by virtue of being explananda in either of three ways: The explananda cannot be the beliefs expressed by the judgements; they cannot be what is reported by considered judgements; nor can they be the making of considered judgements.

If what is reported by considered judgements is taken to be the explanandum, then, as indicated earlier, the distinction between the "explanandum" criterion and the "realism" criterion seems unnecessary. However, I think it is profitable to discuss these two criteria separately, since they raise different, though related issues. In the next section, I consider the view that considered judgements are to be used as assessing standards in virtue of being indicators of moral reality.

## 2. Considered Judgements As Indicators Of Reality

Observation reports might be used as assessing standards because

they are taken to be indicators of non-moral reality. As we pointed out in the above section, scientific theories are to be assessed against the world. Analogously, one could theorize that considered judgements are to be used as assessing standards because what they report are taken to indicate moral reality. One might also try to account for the credibility of considered judgements by appeal to moral realism.

However, there are three general problems with moral theories and moral realism that I wish to discuss in this section, problems that do not obviously arise in connection with non-moral theories and non-moral realism.

We might define moral realism, very broadly, as the view that the truth of moral utterances is to consist in their correspondence with some mind-independent fact or state of affairs. Non-moral realism can be defined in an analogous way.

The first problem concerns the type of moral fact considered judgements are supposed to indicate. With respect to non-moral realism, the corresponding facts are physical or naturalistic. Non-moral theories are assessed against an ontology that is amenable to scientific investigation. With moral realism, on the other hand, it is unclear whether the pertinent moral facts are physical or sui-generis. Theorists who espouse moral realism disagree about the ontology of moral facts. For example, according to Nagel,<sup>11</sup> moral facts are sui-generis, whereas according to Brandt,<sup>11</sup> moral facts are psychological. If a considered judgement is to be used as an assessing standard in virtue of the truth of its utterance corresponding to some type of fact in the world, one ought to know what type of fact is involved. It will not do to claim that a considered judgement is a bona fide assessing standard because the truth of its utterance corresponds to some type of fact but remains unclear as to what type of fact is



involved, other than the (here) unhelpful characterization of it as "moral". For one thing, this would amount to claiming that we are to assess our moral theories against standards whose type is unknown. Or perhaps, more accurately, the type one might say is known: it is moral facts. The point is that this characterization of them is unhelpful given the controversy about what "moral" facts are. For another thing, this mode of reasoning would leave unanswered several pressing questions such as the following: If considered judgements are indicators of moral fact then, presumably, under specified conditions, one should be able to explain the making of a considered judgement by appeal to the relevant moral fact, just as in the non-moral realm, one can explain the making of an observation report by appeal to physical facts. But given the controversy about what "moral" facts are, how can such an explanation be forthcoming? In addition in the non-moral field, there are plausible considerations that substantiate the claim that observation reports, under specified conditions, indicate non-moral reality. However, in the moral field, if one assumes that considered judgements indicate moral facts, and moreover, remains unclear as to whether these facts are psychological, or physiological, or sui-generis, or what have you, then it is difficult to envision how one can provide evidential support for the claim that considered judgements indicate facts. And in the absence of any such evidential support, why accept moral realism?

The second problem that arises with the proposal that considered judgements are to be used as assessing standards in virtue of indicating facts, is that in the absence of arguments for moral realism, this proposal begs the question against non-cognitivist meta-ethical theories.

According to these theories, ethical terms are not fact-stating or property-ascribing terms, and ethical judgements cannot be true or false. To assume, ab initio, that moral judgements are fact-stating, without an argument(s) for this assumption, nor an argument against non-cognitivism is unsound philosophical practice.

And finally, the third problem with the view that considered judgements indicate facts is that it begs the question against theorists like Richard Brandt,<sup>12</sup> and David Copp,<sup>13</sup> who have argued that considered judgements are not credible; i.e., one cannot argue that considered judgements are credible on the basis that they indicate facts without begging the question. Much more will be said regarding considered judgements and their credibility in the next section.

### 3. Considered Judgements And Credibility

A third option is to claim that observation reports are to be used as assessing standards because they are highly credible. "Credible" is one of those "umbrella" terms that can have various interpretations. We could take the proposition "X is credible" to imply that "X is reliable" or "X is not easily defeasible". If we did, an explanation would be owed as to why observation reports are reliable. Causal explanations are generally offered in order to account for their credibility. These causal explanations presuppose non-moral realism. In his appeal to realism, the theorist in science may wish to specify optimal conditions under which observation reports are taken to be indicators of reality. An alternate account as to why observation reports are credible might be offered but what such an account would be is anyone's surmise. I think the most plausible account of the reliability of observation reports is that they

are credible because they are indicators of reality. It is the presupposition that our observation reports give us a "hook on to the world" that gives them their credibility.

Why does the credibility of an observation report or a considered judgement enable either one of them to be used as an assessing standard? If we take considered judgements or observation reports to be credible because they are taken to be indicators of reality, then the answer is clear: They are to be used as assessing standards because they reflect what there really is; our theories are assessed against the world, and not against standards that are arbitrary.

The moral theorist might hold that considered judgements are used as assessing standards because they are credible. Of course, we need to understand how "credible" is to be interpreted in the context of moral theory. If "credible" is to be construed as "reliable", we need to know why they are reliable.

The credibility of considered judgements is crucial to the wide reflective equilibrium theorist, since he assesses moral theory against considered judgements. If considered judgements are not credible, then they cannot properly be used as assessing standards. If they cannot properly be used as assessing standards, then the wide reflective equilibrium program collapses.

As for the credibility of considered judgements, neither Rawls nor Daniels offers an account of why they are reliable, apart from the fact that they are considered judgements. An implausible account of their credibility is that they are credible because under specified circumstances they are self-certifying or self-evident. This account is implausible since

firstly, the notion of self-evidence itself is in need of explanation: If someone claims that it is self-evident that pain is bad, we are not being told anything about the grounds which can be used to substantiate this claim. And secondly, different considered judgements can be self-evident to different people, in which case, it is unclear which considered judgement is to be used as the appropriate assessing standard. A more plausible account of the credibility of considered judgements might be told by invoking moral realism.

Certain difficulties arise for the wide reflective equilibrium theorist if he takes considered judgements to be assessing standards because they are credible. In order to portray what these difficulties are, it will be helpful to discuss Daniels'<sup>14</sup> responses to the "no credibility" objection, and to highlight what is wrong with his responses.

The "no credibility" objection states that unless we have independent reason to believe that our considered judgements in ideal circumstances would themselves be justified, the fact that a morality stands in some justificatory relation to our considered judgements does not show it to be justified. Generalizing, unless it can be shown that a set of considered judgements, or observation reports, under specified relevant circumstances are credible, there is no reason to think that they constitute appropriate standards against which a theory can be assessed.

According to Daniels, the proponent of the "no credibility" objection argues in the following way (p. 270): In the case of observation reports, their credibility can be explained in terms of a causal story. We can show that observation reports are initially credible for some reason other than their coherence within an accepted system of beliefs. However,

there is prima facie reason to think that considered judgements, even sincere ones, are not reliable, and there is prima facie reason to think that no causal story can be told to account for their reliability. Consequently, it is concluded that considered judgements should not be used as assessing standards.

Daniels provides three responses to the "no credibility" argument. We will consider each of them in turn.

(i) In his first response, Daniels claims that the argument rests on an assumed, inappropriate analogy between observation reports and considered judgements. The argument, according to Daniels, presupposes that observation reports and considered judgements function in the same way and on the basis of this, concludes that since no causal story can be told about the reliability of considered judgements, no account of credibility can be given for considered judgements. However, Daniels continues, considered judgements function differently from observation reports - a considered judgement is more like a theoretical than an observational statement.

"The 'no credibility' argument gains its plausibility from the assumption that the analogy to observation reports should hold and then denigrates moral judgements when it is pointed out that they differ from observation reports. If they should and do function differently - because they are different kinds of judgements - that is not something we should hold against the moral judgements." (p.271)

In response to Daniels' response, it must be stressed that the import of the "no credibility" argument is not that the same credibility story must be told for observation reports as for considered judgements, a credibility story that invokes realism. This is how Daniels seems to

be interpreting the "no credibility" objection. Rather, the force of this objection stems from the insistence that some account of the reliability of considered judgements must be forthcoming if these judgements are to be used as assessing standards.

(ii) In his third response, (I will consider the second response shortly) Daniels claims that in the construction of moral theory, there is good reason for starting from considered judgements. Part of his rationale for saying this, I suppose, derives from the observation that although there is disagreement about considered judgements in a culture, there is also agreement.

But here, we should keep in mind the distinction, even if it is only a crude one, between the context of discovery and the context of justification. There might be good reasons for starting with considered judgements in the construction of moral theory but this does not imply that considered judgements should play any role in the justification of a morality. The "no credibility" objection emphasizes that in the absence of an independent credibility account for considered judgements there is good reason not to assess moral theories against considered judgements.

We conclude that Daniels' first and third responses fail to meet the "no credibility" objection. In the next section, we will consider Daniels' second response to the "no credibility" objection. We will argue that this response is also inadequate in that it suggests a dilemma for wide reflective equilibrium theory.

### 3. The Conservative's Dilemma

In connection with Daniels' second response to the "no credibility" objection, I will develop what I will call the "Conservative's dilemma", for conservative wide reflective equilibrium theory.

Daniels' second response to the "no credibility" objection is the most interesting of the three. He suggests that an account of the credibility of our considered judgements will flow or derive from the equilibrium package itself.

The "... 'no credibility' criticism is at best premature. It is plausible to think that only the development of acceptable moral theory in wide reflective equilibrium will enable us to determine what kind of 'fact', if any, is involved in a considered moral judgement." (p.271)

To support this view, Daniels suggests that we are able to assign initial credibility to our observation reports only because an account of their credibility derives from theories in a non-moral wide reflective equilibrium. The intuitive idea underlying Daniels' second response seems to be something like this: In science, we initially assume the credibility of observation reports, perhaps in light of the further assumption that observation reports are indicators of reality. On the basis of these assumptions, we develop scientific theories and assess them, in part, by considering how they cohere or "fit" with the observation reports that are assumed to be credible. The developed theories in turn, will tend to confirm or disconfirm the actual credibility of the observation reports. Mutatis mutandis, a similar account can be given for the assessment of the credibility of considered judgements. This account of how we are to proceed in order to ascertain whether observation reports or

considered judgements are credible has a rather awkward implication, an implication that undermines the plausibility of such an account.

Let me talk about considered judgements first. One wonders what would happen if the theories that are developed on the basis of assuming the credibility of considered judgements were to indicate that considered judgements as a type are not in fact credible. I am envisioning a situation where theories that emerge in an apparent wide reflective equilibrium imply the non-credibility of considered judgements. There are two interesting possibilities here: One could either deny or affirm that one would have a wide reflective equilibrium if the theories in a putative equilibrium package implied the non-credibility of some or all of the considered judgements contained in the package. I will discuss the first possibility in this section and the second possibility in section four.

According to the first possibility one would contend that the theories in the package would not be in equilibrium with the judgements that are non-credible.

If this were the correct way to describe the situation, then some awkward implications would ensue for the theorist: The theorist would be confronted by what we will call the "Conservative's dilemma".

The dilemma would go as follows: First, let the letter J represent a considered judgement held by a person S. Second let the letter P represent a package of theories held by the same person. Suppose that J is not coherent with the package P, so that the person in question is not in wide reflective equilibrium.



The first horn of the dilemma: Now either J impugns as unjustified the package P or it does not. If, on the one hand, J does not impugn as unjustified the package P, then wide reflective equilibrium theory cannot be correct. This is because wide reflective equilibrium theory asserts that a set of moral theories is justified only if it is in wide reflective equilibrium with considered judgements. But in our case, the package P is not in wide reflective equilibrium with the person's considered judgements. Hence if wide reflective equilibrium theory is to be maintained, the existence of the considered judgement J must be taken to show that S is not in wide reflective equilibrium and that the moral theories held by S are not justified.

The second horn: Suppose, on the other hand, J impugns as unjustified the package P. If it does, then the wide reflective equilibrium theorist faces the following problem. Again the problem is in the form of a dilemma. On the one hand, the theorist could hold that J is not justified. However, if a theorist held that J is not justified, then it would be difficult to understand how the existence of J among the person's beliefs, even if J is an intractable judgement, is sufficient to show to be unjustified the package of theories in question. How can a belief which is not itself justified show to be unjustified a whole set of theories held by a person, and held perhaps with as much conviction as the belief in question. It seems that the theorist must allow that J is at least credible, even if it is not itself justified, in order to make plausible the idea that J, among the beliefs held by S, is sufficient to show P to be unjustified. But this leaves a problem for wide reflective equilibrium theory, for how can it be that a judgement is credible even

while it is not in wide reflective equilibrium with a set of theories held by the person in question? As David Copp<sup>15</sup> indicates, the best a wide reflective equilibrium theorist could do, at this point, would be to say that the mere confidence with which J is believed accords to J a kind of credibility sufficient to enable J to undermine the package of theories in question. However, in his paper, professor Copp points out that the theorist seems to have no ground for taking such a position.

On the other hand the theorist could hold that J is itself justified. Were J an intractable judgement, he might be especially motivated to do so. However, if he did, then he would have abandoned wide reflective equilibrium theory because wide reflective equilibrium theory asserts that a moral judgement is justified only if it is in wide reflective equilibrium with a set of moral theories, and in this case, J is not in wide reflective equilibrium.

Now, the theorist might try the following reply:<sup>16</sup> "Either J would be among S's considered moral judgements in wide reflective equilibrium or it would not. If it would be, then it is capable of impugning the package P as unjustified but then too, it is justified. On the other hand, if it would not be, then it is not justified and it is not capable of impugning P as unjustified. This is because P might turn out to be the very package of theories that would be held by S in wide reflective equilibrium. The theorist might propose that in this way he is able to escape the dilemma. However, a third possibility exists. That is, it could be that S is psychologically incapable of reaching a wide reflective equilibrium, and that he would remain attached to J no matter what. In that case, the fact that S holds J as a considered judgement would be

sufficient to show that the package P is not justified, because it would be sufficient to show that P would not be held by the person in wide reflective equilibrium. Moreover, J itself is unjustified because J would not be among the considered judgements that S would hold in wide reflective equilibrium. There are no judgements that S would hold in wide reflective equilibrium. Hence, the theorist is left with the problem that J is able to impugn as unjustified a package of theories even though J is not itself justified.

#### 4. A Second Attempt To Circumvent The Conservative's Dilemma

Towards the end of section (2), we were discussing Daniels' suggestion that an account of the credibility of considered judgements will derive from a wide reflective equilibrium. In evaluating this suggestion, we considered the situation where theories purported to be in a wide reflective equilibrium implied the non-credibility of considered judgements. We said that with regard to this situation, there were two possibilities: First, one could deny that one would have an equilibrium if the theories in a given putative equilibrium package implied the non-credibility of some of the considered judgements contained in the package. This possibility led to the discussion of the case where we have a considered judgement outside of reflective equilibrium impugning a package of theories as unjustified. This discussion in turn, led to the development of the "Conservative's dilemma".

Let us now consider the second possibility. According to the second possibility, in a situation where the emerging theories implied the non-credibility of some of the judgements with which the theories are supposed to cohere, one could acknowledge that the theories would be in an

equilibrium with these considered judgements. After all, if I have interpreted Daniels' third response to the "no credibility" objection correctly, what the theorist is supposed to do is to assume the credibility of considered judgements as a type. On the basis of this assumption, we are to construct and assess moral theory. That is, we are to proceed with the assumption that considered judgements are credible and work our way to an equilibrium. Prima facie, there does not seem to be anything incoherent in the idea that some of the considered judgements contained in an equilibrium package, that have been assumed to be credible, are in actuality not credible. In fact, this seems to be the most plausible way of interpreting Daniels' claim that a credibility story, or more accurately, an account of the credibility status, of considered judgements, will flow or derive from a wide reflective equilibrium: Any theorist must initially allow the possibility that considered judgements are not credible. But if considered judgements are not in fact credible, then the first interpretation or possibility viz, that one would not have an equilibrium if the theories in a given putative equilibrium package implied the non-credibility of some of the considered judgements contained in the package, denies that a story about the credibility status of considered judgements can derive from a wide reflective equilibrium. This is so since according to the first possibility, if the theories in a given putative equilibrium package implied the non-credibility of the considered judgements contained in the package, then one would not in fact have theories in an equilibrium. It would be theories in a putative equilibrium package that implied the non-credibility of considered judgements. But according to the second possibility, one would have theories in an actual

equilibrium package, not merely a putative equilibrium package, that implied the non-credibility of the considered judgements in that package.

This second possibility, perhaps, would enable a theorist to circumvent the "Conservative's dilemma", since, according to this possibility, it would not be the case that one would have a considered judgement outside of an equilibrium impugning a package of theories as unjustified.

But if it is a possibility that considered judgements contained within an equilibrium package are not in fact credible, then this second possibility also promises untenable implications: First, according to this second possibility, a theory may be in equilibrium, or may be coherent, with a considered judgement even though the theory implies that the judgement is not credible. This is a rather unusual reading of "coherent". At least, this interpretation of "coherent" differs radically from the interpretation suggested by the R/D theory. According to the R/D theory, equilibrium implies an harmonious relation between background theories, principles and considered judgements.

And second, if the theories contained within the equilibrium package itself imply the non-credibility of any of our considered judgements, then if considered judgements play an ineliminable role in (i) constraining background level III theories and (ii) assessing the theories themselves, the wide reflective equilibrium program must be abandoned. It must be abandoned since, firstly, the theories in wide reflective equilibrium have been developed on the basis of non-credible data. If these theories have been developed, even partly, on the basis of non-credible data, we have no reason to suppose that they

are tenable. And secondly, these theories have been assessed against standards that are not credible. Call this result the "paradoxical result".

If the emerging theories do in fact imply the non-credibility of our considered judgements, what one might try to do in order to save wide reflective equilibrium theory is (i) alter some of the background theories, or (ii) start with an alternate set of considered judgements, and hope that this time around, the emerging theories will yield a more amenable answer, that is, that considered judgements are in fact credible.

However, if a wide reflective equilibrium theorist opts for revising the background theories, then he is simply rigging theory in order to obtain a desired result. Why would revision of the background theories imply rigging? Consider, first, an analogy with Physics. Suppose, a theory is desired to show that Earth's orbital path around the sun is circular. Suppose, however, the best physical theories indicate that Earth moves in an elliptical path around the sun. Suppose, further, that this result is undesirable and physicists resort to revising physical theory in order to obtain the desired result. Given that the best theories indicate that Earth's orbital path is elliptical, any revision to show otherwise would presumably involve the use of ad hoc hypotheses, and the alteration of well-embedded physical principles. In addition, it is also unclear whether such revision would be possible; it is unclear, for example, whether such revision would be compatible with observation reports. Clearly, any such change to physical theory would be a case of rigging.

The situation is similar with regard to altering background theories in order to show that considered judgements are credible: Suppose we have been working towards an equilibrium, and we have been doing so for a considerable time. We have made a number of "false" starts, but eventually we have reason to believe that the moral theories we are working with are pretty tenable. Presumably, the moral theories that are to cohere with principles and considered judgements in wide reflective equilibrium are going to be theories that are deemed to be the most plausible. These theories, in turn, are going to be constrained by empirical theories which, presumably, are also the most tenable empirical theories. Suppose these theories implied that considered judgements are not credible. Given that these theories are the most plausible theories, were one to alter them in order to show that considered judgements are credible, this would again seem to be a case of rigging. Successful revision is possible, but revision might involve renouncing well established physical principles, and credible philosophical principles. In short, option (i) is not open to the conservative.

According to option (ii) the conservative is to start with an alternate set of considered judgements, work towards an equilibrium, and hope that this time around, theories in equilibrium imply the credibility of considered judgements. But this option also has its shortcomings. First, it is unclear how a theorist is to literally start with an alternate set of considered judgements. Considered judgements are not the kind of things that people change at random. This is partly due to the fact that the considered judgements that one holds are a function of one's psychological and sociological constitution. And it is evident that one cannot switch one's psychology, or one's sociological constitution, at

random. Second, considered judgements constrain the development of background moral theories. It is plausible to assume that different sets of considered judgements will lead to the development of different types of background theories. One type of theorist, for example, might prefer a theory of the person that is non-egalitarian, whereas somebody with a Rawlsian bent might prefer a theory of the person that is egalitarian. Suppose the best theories in equilibrium implied the non-credibility of considered judgements. Then, one would be reluctant to start with an alternate set of considered judgements, in light of the fact that an alternate set of considered judgements would probably lead to the development of radically different background theories of personhood, moral acquisition, etc. In other words, adoption of an alternate set of considered judgements, will presumably affect a number of theories, and not simply the particular theory that implies the non-credibility of considered judgements. And with regard to these other background theories, why alter them, via the adoption of an alternate set of considered judgements, if these theories are highly plausible?

Third, the best background theories might imply that considered judgements as a kind are not credible, in which case any set of considered judgements would be unsuitable candidates for assessing standards.

There is one final consideration that we wish to discuss in relation to Daniels' claim that an account of the credibility of considered judgements will derive from a moral equilibrium. It appears that Daniels toys with this interesting idea because he thinks that:

"Properly understood, the credibility story about non-moral observation reports is itself only the product of a non-moral wide reflective equilibrium of relatively recent vintage."  
(p.272)



I wish to argue that there is something fundamentally suspect about the claim that a credibility story for observation reports derives from a non-moral wide reflective equilibrium. If my arguments to this effect are cogent, then we have one less reason to think that an account of the credibility status of considered judgements will derive from a moral equilibrium.

First, recall that in order for theories to be in wide reflective equilibrium, or in order for them to be coherent with each other, it is not sufficient that the theories merely be consistent with each other - the theories must stand in some complex coherence relation R. The connecting links between elements in a coherence system, for example, might be one of inference, or of explanation, or of mutual confirmation.

In claiming that a non-moral wide reflective equilibrium package exists, one must be claiming that non-moral theories stand to each other in a complex coherence relation R. But surely this is contentious. It might be true that theories in physics, or in sociology, or in biology, or in mathematics, stand to each other in relationship R, but it would be much more difficult to establish the claim that sociological theories and physical theories stand to each other in relation R. This is, of course, not to deny that sociological theories and physical theories may well be consistent. It appears that talk about one huge non-moral wide reflective equilibrium might be misleading. Unless we espouse some kind of reductionism, for example, that all true theories are reducible to basic physical theory, little sense is to be made of the claim that neurophysiological theory is contained within the same equilibrium package as quantum theory. This again, is not to deny that the principles

of neurophysiological theory will probably be consistent with those of quantum theory. In science, it might well be the case that there are several equilibrium packages, each peculiar to its own region or area of science. But were this true, it would be misleading to claim that a credibility story for observation reports derives from a single non-moral wide reflective equilibrium.

Furthermore, it is entirely possible that certain philosophical theories and theories of perception that imply the credibility, or non-credibility, of observation reports are not part of any equilibrium package. But if this state of affairs is possible, then it would be false to claim that a credibility story for observation reports derives from a non-moral equilibrium.

To account for these problems, a conservative like Daniels might weaken the notion of equilibrium. He might propose, for example, that to say that an account of the credibility status of considered judgements flows from an equilibrium is just to say that an account of their credibility flows from theories that we believe. This is true, but it does not tell us anything that we do not already know: Presumably, the credibility status of considered judgements is to be established by argument and theory. Or, he might propose that to say that an account of the credibility status of considered judgements flows from an equilibrium is just to say that an account of their credibility flows from a set of theories that are consistent. But this fact reduces the notion of coherence to mere consistency, and most coherence theorists, Daniels included (p. 258) eschew this.

Finally, since Daniels wishes to claim that a credibility story for

observation reports derives from a non-moral wide reflective equilibrium, there is a final consideration that he must take account of.

We should distinguish between internal theories, theories that are contained within an equilibrium package P, which cohere with each other, and external theories, theories that are not contained within equilibrium package P, but which are consistent with the internal theories of P. It must be remembered that mere consistency is not a sufficient condition for coherence (see chapter 2); so even if the internal and external theories are consistent, this does not imply that they cohere with each other in any strong sense of "coherence", and consequently, it does not imply that they are contained within the same equilibrium package. This point might be difficult to grasp in the context of scientific theory but perhaps we might be able to understand it better if it is made in connection with moral theory. Surely, Daniels must wish to distinguish between those theories that are contained within an equilibrium package  $P_1$  and that justify a given morality, and those theories that are consistent with the theories in  $P_1$ , but are not themselves contained within  $P_1$ . In a similar way, there must be a criterion (or a set of criteria) that enables us to distinguish between those internal theories that are contained within a non-moral wide reflective equilibrium package M, whatever this is, and however we are to circumscribe such a package, and those external theories that are consistent with theories in M, but which are not contained within M.

One might balk at the idea that one is able to individuate equilibrium packages, but the coherence theorist, in any case, seems committed to being able to differentiate different cognitive systems.

As mentioned in Chapter 2, according to the coherence theorist, for a system of beliefs to be justified, the system must be more coherent than any other currently available alternative.

Let us assume that the most plausible account of the credibility of observation reports is an account in terms of a causal explanation. Then in claiming that a credibility story for observation reports derives from a non-moral wide reflective equilibrium, Daniels must be claiming that a causal account of the credibility of observation reports itself derives from a non-moral wide reflective equilibrium. If we distinguish, and I am claiming that we do, between those internal theories that are contained within an equilibrium package and those external theories that are consistent with the internal theories but not contained within the same equilibrium package, then there must be a "delimiting" criterion that enables us to distinguish between internal and external theories. In order for Daniels to claim that a causal theory of the credibility of observation reports is an internal theory, which he must wish to claim, if he intends to hold that a credibility story for observation reports derives from a non-moral wide reflective equilibrium, he must show that such a theory is an internal theory by appeal to a delimiting criterion. But neither has he formulated such a criterion, nor has he shown that causal theory is an internal theory.

To tie a few ends together, I have suggested, in the last few sections, that considered judgements can serve as assessing standards if they are credible. Daniels, in his discussion of the "no credibility" objection acknowledges this. Daniels' first and third responses to the "no credibility" objection are clearly inadequate. As far as Daniels'

second response is concerned, I offered two interpretations of the suggestion that an account of the credibility status of considered judgements will flow or derive from wide reflective equilibrium itself. Each interpretation has undesirable implications. I conclude that Daniels has not refuted the "no credibility" objection. The inability to refute the "no credibility" objection suggests that Daniels' theory of justification provides a derivative test for the justification of a morality. It provides a derivative test, since, the very standards against which a morality is to be assessed and hence justified, are themselves in need of independent or prior justification.

#### 5. The Revisionist's Dilemma

Before we leave this chapter, I wish to consider a difficulty which arises specifically in relation to a theorist who utilizes a revisionist model of wide reflective equilibrium. This difficulty suggests another dilemma for revisionist wide reflective equilibrium theories. We said earlier that type 1 revisionists permit extensive revision of both the initial set of considered judgements and considered judgements held in wide reflective equilibrium, whereas type 2 revisionists restrict revision to the initial set of considered judgements. These revisions may be theory-based or they may not be so based.

Let us assume that theoretical investigation has established that considered judgements are credible. Let us also suppose that the theory (or the theories) establishing credibility specifies a relevant condition (or conditions) under which our considered judgements are credible. Now either considered judgements are credible if and only if in equilibrium, or it is not the case that they are so credible.

On the one hand, let us suppose that the relevant condition is not the condition that considered judgements are credible if and only if in wide reflective equilibrium. The theory establishing credibility, for example, permits the state of affairs where considered judgements are credible prior to equilibrium. Given these assumptions, if a revisionist like Daniels, permits extensive revision of our considered judgements, even theory-based revision, in order to reach equilibrium, then this amounts to bad theoretical practice, since, if considered judgements are credible prior to equilibrium, and if we reject a set of them in order to reach equilibrium, we are rejecting credible data.

The revisionist might try the following reply. Suppose the relevant condition is a matter of confidence and theoretical support. Suppose, furthermore, that the very confidence with which a considered judgement is believed imputes to it a kind of credibility sufficient to enable it to impugn theories not in wide reflective equilibrium with it. Suppose however that lack of theoretical support undermines our confidence in a given judgement J, and so we reject it. However, when we reject it, J is not a considered judgement, for we have lost confidence in it, and it is not credible, for we have assumed that credibility is a function of confidence. Consequently, when we reject J, we are not rejecting credible data, let alone rejecting a considered judgement. The theorist might propose that in this way, he can escape the first horn of the dilemma.

However, this reply is inadequate. Its principal shortcoming resides in the premise that the credibility of considered judgements is a function of the confidence with which a judgement is believed.

There are considerations which seem to render this premise implausible: Suppose, unknown to him, a person in early childhood had literally been conditioned to believe that abortion is wrong. In adulthood, this person believes the judgement, J, "Abortion is wrong", with great confidence, and he may even adduce theoretical considerations to justify his belief in J. However, given the way in which he acquired his confidence in J, it is unreasonable to claim that J is credible merely because it is believed with great confidence. If the credibility of a judgement is primarily a matter of the confidence with which the judgement is believed, then judgements like J which we would not deem to be credible, would in fact turn out to be credible. The confidence with which a judgement is believed is not a sufficient condition for the credibility of the judgement. But is it even a necessary condition? Is it not possible for a judgement to be credible even though we do not believe it with any great confidence? Surely, "Yes." Given that the confidence with which we believe certain judgements varies with factors such as one's childhood experiences, one's social environment and one's education, it does not even seem that the confidence with which a judgement is believed is a necessary condition for the credibility of that judgement.

However, the theorist has a second reply that might allow him to escape the charge of bad theoretical practice in virtue of rejecting credible data. He might say that there is a difference between rejection and revision. He might argue that revisions are theory based and so legitimate. One gradually works through and loses (or gains) confidence in a particular judgement. To meet this response, we can argue in the following way: Revision can work in both directions. Not only can

judgements be revised in order to "match" with theory but theory can also be adjusted. Presumably, if the judgements are extensively revisible, then even though credible, they must be credible to only a low degree, since, if they were highly credible, we would be reluctant to revise them extensively. As Rawls<sup>17</sup> points out in connection with science,

"... if we have an accurate account of the motions of the heavenly bodies that we do not find appealing, we cannot alter these motions to conform to a more attractive theory."  
(p.49)

Revision of highly credible data, even if theory based, is bad theoretical practice.

If it is a necessary condition that considered judgements be minimally credible for extensive revision of them to be permitted, then the revisionist is assessing theories in putative or partial wide reflective equilibrium against standards that are easily defeasible. But if this is so, what extra justificatory force can be obtained by the successful testing of a theory against standards that are easily defeasible? Here a revisionist might rejoin that most of the justificatory force in wide reflective equilibrium derives from the level III and IV background theories (see chapter 3) and not from level I partial equilibrium. But if this is the case, then there is even more reason to wonder why considered judgements at level I are included in the justificatory process for theories in wide reflective equilibrium.

A revisionist might reply that even if it is a necessary condition that considered judgements be credible only to a low degree for extensive revision of them to be permitted, he need not be committed to the view



that he is assessing theories in wide reflective equilibrium against easily defeasible standards. He need not be committed to this view since the relevant condition under which considered judgements are credible, might coincide, even necessarily coincide, with the property of being subscribed to confidently in wide reflective equilibrium.

Let us suppose, then, on the other hand, that the credibility of considered judgements is established by a theory which specifies that considered judgements are credible only in wide reflective equilibrium. It is only our considered judgements in wide reflective equilibrium that are to be used as assessing standards.

But if we are to grant this, then it is difficult to understand how considered judgements can be used as assessing standards: It seems that it is an essential characteristic of an assessing standard that such a standard should allow for the possibility that a theory which is to be assessed against such a standard be false, or inadequate, or "incorrect" for the reason that the theory does not "stand up" to the required standard. For example, in the non-moral realm, an observation report might be treated as an assessing standard because what is reported by the observation is the explanandum. And we might claim that a non-moral theory is inadequate because the theory conjoined with appropriate bridge laws does not imply the explanandum. However, if a considered judgement in wide reflective equilibrium is to be used as an assessing standard, then it can never be the case that the theories in wide reflective equilibrium which are to be assessed against this judgement be unjustified. This is because when wide reflective equilibrium is reached, by definition, theories in wide reflective equilibrium cohere

or "match" adequately with the considered judgements in wide reflective equilibrium.

Summarizing this chapter, we have been concerned to understand why considered judgements are to be used or treated as assessing standards in wide reflective equilibrium theory. I suggested, first, that they may be so used if they are taken to constitute the explananda. I offered three interpretations of this suggestion and argued that none of them are plausible. I next considered the proposal that considered judgements are to be treated as assessing standards in virtue of being indicators of reality. This proposal also has its shortcomings. Finally, in the last half of this chapter, I have suggested that considered judgements in conservative theories might be used as assessing standards because they are credible. I have shown that the "no credibility" objection has not been met, and this in turn suggests that conservative wide reflective equilibrium provides a derivative test for the justification of moral judgements, moralities, or theories of moral justification. In addition, I have argued that if a wide reflective equilibrium theorist wishes to permit extensive revision of the initial set of considered judgements, then to avoid the charge of bad theoretical practice, he must assume that considered judgements are credible only to a low degree. But if he does this, the rationale of using considered judgements as assessing standards is undermined.

REFERENCES

1. Rawls, John, A Theory of Justice, Cambridge: Harvard University Press, 1971.
2. Daniels, Norman, "Wide Reflective Equilibrium and Theory Acceptance in Ethics", The Journal of Philosophy, 76, 1979, pp. 256-282.
3. Hempel, C.G., Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, New York: The Free Press, 1965.
4. See: Brandt, Richard B., Ethical Theory, Englewood Cliffs, N.J.: Prentice Hall, 1959, pp. 173-176.
5. Firth, Roderick, "Ethical Absolutism and the Ideal Observer", Philosophy and Phenomenological Research, 12, 1952, pp. 317-345.
6. Harman, Gilbert, The Nature of Morality, New York: Oxford University Press, chapter 1, 1977.
7. Ibid.
8. Ibid.
9. Ibid.
10. Nagel, Thomas, "The Objectivity of Ethics", unpublished, 1980.
11. Brandt, Richard B., A Theory of The Good and The Right, Oxford: Clarendon Press, 1979.
12. Ibid, chapter 1.
13. Copp, David, "Considered Judgements and Moral Justification: Conservatism In Moral Theory", forthcoming in David Copp and David Zimmerman, Morality, Reason and Truth, Dordrecht, Holland: Reidel, 1982.
14. Daniels, Norman, Ibid, 1979.
15. Copp, David, Ibid, 1982.
16. I am indebted to David Copp for the contents of the next paragraph, and for helpful discussion on the "Conservative's dilemma".
17. Rawls, John, *ibid*, 1971.

## Chapter 5

### OBJECTIVE AND SUBJECTIVE JUSTIFICATION

It is important to keep in mind, as David Copp<sup>1</sup> has stressed, the distinction between theories as to when a morality itself is justified and theories as to when a person is justified in holding a morality or a moral judgement. In this chapter we will first discuss this distinction. We will then propound considerations that suggest that it is the former notion that is germane to theories of moral justification. Finally, we will argue that the R/D theory is best construed as a theory specifying conditions under which a person is justified in believing a morality or a moral judgement, and, consequently, it is not a theory relevant to the justification of a morality itself.

#### 1. Objective and Subjective Justification

As a first attempt to elucidate the distinction between theories of moral justification, and theories of the circumstances under which a person is justified in holding a morality, we can think of circumstances in which a person would be justified in his beliefs but where the beliefs themselves could well be unjustified. For example, a person brought up in a religious community, without any critical exposure to religious doctrine may be justified in his belief that there is a God, but the belief that there is a God may not itself be justified. To take another example, on the basis of the theories that he has developed, a scientist may be justified in his belief that there is life on Jupiter but this belief itself may well be unjustified. As a third example, suppose Jones is brought up in a society that condones apartheid. By simply "following in the shoes" of his family and peers, Jones acquires non-temporary and sincere confidence in the

precept that apartheid is right. Here again, we can claim that perhaps Jones is justified in his belief that apartheid is right. After all, Jones, like a great many of us, has simply accepted the values inculcated in his culture. But surely Jones' being justified in accepting the values of his culture does not entail that the values themselves are justified.

In describing conditions under which a person is justified in his views, both moral and non-moral, an appeal is frequently made to psychological and sociological theories. A sociologist, for example, after studying the customs of a society, might be in a good position to explain why the members of a society are justified in their belief that cannibalism is right. But in developing a theory that explains such behaviour, he will not have developed a theory that justifies the judgement or the proposition that cannibalism is right. Similarly, a psychologist might be in a good position to explain how a person arrived in his belief that the subjugation of women is right. And in doing so, the psychologist might be able to explain why the person is justified in his belief that the subjugation of women is right. But again, such an explanation would not imply anything about whether the proposition that the subjugation of women is right is itself justified.

In developing theories that specify conditions under which a moral judgement is itself justified, either simpliciter or in relation to a person or a group, it is evident that we are not interested in how a person or group comes to acquire the moral beliefs that they subscribe to. Rather, it is the justificatory status of the moral judgements or the beliefs themselves that we are interested in. In developing theories that explain how persons come to believe what they believe, we might learn something very important about the creative process or "the context of discovery". But it

is clear that in developing theories that purport to justify moral or non-moral propositions, it is not the context of discovery that we are interested in; we might say that here it is the "context of justification" that is of primary importance.

There seem to be at least three ways to specify, with greater clarity, and precision, the distinction between (a) situations in which a person is justified in believing a moral judgement, and (b) situations in which the moral judgement believed by a person is itself justified. Let us call the type of justification involved in (a) and (b) "subjective" and "objective" justification respectively.

(i) When it is claimed that a moral judgement P is not objectively justified but it is subjectively justified vis a vis subject S, there is the suggestion that P is not objectively justified because somehow, and somewhere, in connection with judgement P, S has made a mistake. First, suppose that S believes that abortion is wrong, and he believes this on false grounds. However, S fails to know that the grounds on which his belief is based are false. Here, S may be justified in believing that P, although P itself is not objectively justified. P is not objectively justified, in part, because it is believed on false grounds. I think we would want to hold that it is at least a necessary condition for P's being objectively justified that belief in P is based on true grounds.

It must be stressed that not all failures of objective justification leave enough to enable us to say the believer is subjectively justified. For example, we might want to distinguish between the following two cases: Firstly, Jones believes that abortion is wrong on false and unreliable grounds. Here we would want to say that Jones is neither subjectively nor objectively justified in believing P. Secondly, Jones believes that

abortion is wrong on false but reliable grounds. Here we may want to allow that Jones is subjectively justified in believing P.

In other cases where P is subjectively but not objectively justified, we suggest that one reason why P may not be objectively justified is that relative to some more objective stance or relative to some higher standard of rationality, the individual is in error. Since the individual is in error, this second explication is a variant of the first. Here we would want to claim that since subject S is not being objective enough, P is not itself justified in relation to S. A problem with this suggestion is in understanding just what is implied by "more objective stance" or "higher standard of rationality." How can objectivity be acquired? Or, under what conditions is objectivity disparaged? I do not pretend to answer, or even be able to answer, these questions to any extent that will do justice to them, but I will venture a few suggestions.

First, when we move from subjective to objective justification, we transcend personal idiosyncracies. One reason why S may not be being objective enough, in relation to belief P, is that S may believe P, or may come to believe P on the basis of partial considerations. The partiality may be relative to the subject himself, or it may be relative to a larger group. In the former case, S may believe that the subjugation of women is right on the basis of personal experience. In the latter case, S may believe that the subjugation of women is right because this moral precept is part of the morality of the group that S belongs to.

In cases of partiality, whether or not S is subjectively justified in believing P might be a function of whether the partiality is relative to S's society as a whole, or relative to S himself. If the partiality derives from rules that are embedded and subscribed to in S's society, then here we might

allow that S is subjectively justified in believing that P.

Secondly, S can fail to be objective enough in virtue of failing to consider relevant and adequate information that bears on the issue in question. Here we would want to say that had S been in possession of this information, he would not have been justified in his belief that P. What is to count as relevant and adequate information will directly influence the objectivity that one is able to achieve. If, for example, what is to count as relevant information is the science of the day, then relative to this information, one would be being more objective were one to take relevant information to be constituted by the results of an ideal science or the results of science "in the limit".

Again, we want to hold that not all failures of objective justification arising from lack of adequate information qualify as instances of subjective justification. For example, in the event that more information could have been obtained relatively easily and without significant cost to the agent, we would be inclined to say that the agent is not subjectively justified in believing that P.

Finally, to complete this very brief discussion on objectivity, we can end with Thomas Nagel's<sup>2</sup> extremely interesting suggestion that:

"Objectivity is a method of understanding.....To acquire a more objective understanding of some aspect of the world, we step back from our view of it and form a new conception which has that view and its relation to the world as its object. In other words, we place ourselves in the world that is to be understood. The old view then comes to be regarded as an appearance, more subjective than the new view, and correctable or confirmable by reference to it. The process can be repeated, yielding a still more objective conception" (p.1).



Nagel seems to agree that objectivity is to be gained by transcending one's personal conception of oneself and the world - that relative to this objective stance, one's former views can be regarded as subjective. In addition, according to Nagel, one's subjective views are correctable by reference to one's objective views. This suggests that Nagel would agree with the notion that a judgement may fail to be objectively justified vis a vis a person because in connection with this judgement, the person has made an error. That he has made an error would be substantiated if this person were to "move" to a more objective stance.

(ii) Secondly, in at least some cases where a person is justified in believing P, but where P itself may not be justified, we suggest that P may not be objectively justified because of the way in which the belief in P is acquired. To illustrate, suppose S, in childhood, is literally conditioned to believe that abortion is wrong, and as a result of this conditioning, he comes to acquire non-temporary and sincere confidence in this belief. Given that S literally acquires his belief that P via conditioning, S is justified in his belief that P, but this leaves open the question of whether P is itself justified. Here we are inclined to say that S is justified in his belief that P because it would be very difficult for S to rid himself of this belief. The belief and the associated motivational underpinnings have become part of S's "permanent" psychology. To take another example, suppose S's father is an alcoholic. As a result of his father's drinking habits, S has an extremely unhappy childhood. S gradually comes to have sincere and non-temporary confidence in the judgement that any adult male who touches alcohol is morally depraved. Here again, it seems we have grounds for claiming that S is justified

in his belief but the belief is not itself justified. It is difficult to specify principles that distinguish "legitimate" and "illegitimate" means or ways of acquiring beliefs. But even if we fail to specify any such principles, we can still rely on paradigm examples that illustrate what we have in mind when we claim that a belief has been acquired in an "illegitimate" way. Or more to the point, paradigm examples help us to understand why the manner in which certain beliefs are acquired inclines us towards the view that the beliefs in question may not be objectively justified.

In summary, we have suggested that a belief P may be subjectively justified in relation to subject S but not objectively justified when at least three conditions are true:

- (i) S's belief in P is based on some kind of error;
- (ii) S's belief in P is based on S's failing to be objective enough, and
- (iii) S's belief in P is acquired in an "incorrect" or "illegitimate" way.

We stress that we are not involved in formulating necessary and sufficient conditions for a belief being justified itself, and for a person being justified in his belief. We simply wish to make credible that there is this distinction, and to illustrate cases relevant to each of the notions in question.

I think it will be helpful, in further understanding the distinction between objective and subjective justification, briefly to consider Richard Brandt's<sup>3</sup> theory of moral justification. Apparently the theory purports to be a theory of objective justification. However, certain

features of the theory, particularly features related to the way in which certain moral beliefs are acquired, tend to suggest that the theory is really describing conditions under which persons are justified in holding moral beliefs.

## 2. Richard Brandt's Reductionism

Brandt's theory of morality is a replacement or eliminative theory: The theory proposes to redefine traditional questions of morality in terms of rationality defined in a particular way. What makes something right or wrong ultimately rests on the attitudes, desires and aversions a person would have if he were fully rational. Since rationality is a psychological notion for Brandt, moral claims are reduced to psychological claims.

A person's action is fully rational in Brandt's sense of the word if and only if two conditions are met:

- (i) Any desires affecting a person's tendency to perform the act are rational. These desires are rational only if they survive cognitive psychotherapy, i.e. a

"...process of confronting desires with relevant information, by repeatedly representing it, in an ideally vivid way, and at an appropriate time...."  
(p.113)

and,

- (ii) the person would still perform the act if he were vividly and adequately to represent to himself all relevant, available information.

"Relevant information" means all information that would affect a person's decision to perform a certain act. (pp. 12, 113)

By "available information", Brandt means facts or truths that are accepted by the science of the day and that are publicly known.

(pp. 13, 112)

With this background, Brandt proposes that the question "What is the best thing to do from an agent's own point of view?" be redefined or reformulated as "What would an agent do if he were fully rational?"

From his initial analysis of rationality, Brandt then proceeds to derive reforming definitions for "morally right" and "morally wrong". Moral rightness and wrongness are defined in terms of a person's or a society's moral code. A moral code consists of a set of motivational constraints on behaviour. A social moral code exists when the dispositions to act or the motivational constraints on behaviour are commonly shared by all members of a society.

Next, Brandt argues that a "correct" or justified moral code is a code that is criticized by facts and logic as far as possible. (p.185) Since a code supported by a fully rational person would be a code that has been criticized by facts and logic as far as possible, identifying a code that a fully rational agent would support is tantamount to justifying that code for the agent. (pp. 185, 186) In a similar way, identifying a code that all fully rational agents of a society would tend to support, is tantamount to justifying a social moral code. The notion of supporting a code has to do with advocating the code, teaching it to others, using it as a basis for making long range decisions etc. (pp. 188, 193) A fully rational person viz, a person whose desires are

rational and who would act on these desires if he were to represent to himself all relevant, available information, would support a code if its motivational constraints on behaviour were rational and if he would still act on these rational motivations when he were to vividly and adequately represent to himself all relevant, available information.

Brandt suggests the following reforming definition for "morally wrong":

"'Is morally wrong' (is assigned the descriptive meaning) 'would be prohibited by any moral code which all fully rational persons would tend to support, in preference to all others or to none at all, for the society of the agent, if they expected to spend a life-time in the society'" (p.194)

If it were to happen that all fully rational agents did not support the same moral code for their society, Brandt suggests that we should define "morally wrong" in terms of an individual's moral code - an act is wrong if it is prohibited by a fully rational agent's moral code.

It might appear that the condition of "full rationality" enables Brandt to escape the charge that his theory delineates conditions under which a person is justified in his moral views rather than conditions under which a morality is itself justified. This might appear so, since, as we indicated above, one of the conditions under which it can be claimed that a moral judgement may not be objectively justified, is the condition that an agent fails to be objective enough, or rational enough. The condition of "full rationality" in Brandt's theory suggests maximal rationality.

However, I think it is a mistake to construe Brandt's theory as a theory of objective justification. His theory of rationality can allow

for situations where one would be inclined to claim that certain moral beliefs have been "illegitimately" acquired. Since these beliefs or judgements have been "illegitimately" acquired, a person might be justified in believing these judgements, but the judgements themselves may not be justified. It is quite possible, for example, that even a fully rational agent, due to traumatic childhood experiences, would support a code that condoned the subjugation of women.<sup>4</sup> The judgement "Subjugation of women is right" would then be justified for this agent. But one would want to claim that given that this agent, though fully rational in Brandt's sense of the term "rational", was subject to traumatic experiences in his early childhood, he is justified in his belief that the subjugation of women is right. It would not follow, however, that this judgement is itself justified.

In summary, it seems more appropriate to construe Brandt's theory as describing or specifying conditions under which fully rational agents would be justified in their moral views, rather than specifying conditions under which the moral views themselves would be justified for such agents.

### 3. The Rawls/Daniels Theory As A Theory Of Subjective Justification

I now wish to argue that the R/D conservative wide reflective equilibrium theory is best construed as a theory that specifies conditions under which persons are justified in their moral views, rather than a theory that justifies moralities themselves.

The R/D theory purports to be a relational theory of moral justification. According to this theory, a moral judgement is justified in relation to a person if it is the one that that person would hold

in wide reflective equilibrium. At first blush, it appears that the R/D theory is a theory that justifies moral judgements or moral propositions.

However, a distinctive characteristic of the R/D conservative theory is the ability of considered judgements to impugn the constituents that might be contained in an equilibrium package. In other words, conservative theorists like Daniels and Rawls wish to use considered judgements, under specified conditions, as standards against which moralities are to be assessed. In so using considered judgements, these judgements are accorded a privileged status over moral judgements that are not considered. It is this property that considered judgements instance, the property of being considered, that enables such judgements to impugn constituents of equilibrium packages as unjustified, and, the role that such judgements play in the attainment of equilibrium, that makes credible the claim that the R/D theory is best construed as a theory describing conditions under which a person is justified in his moral views. This requires explanation:

Considered judgements are known to vary extensively - inter societal and intra societal variation of considered judgements elicited in response to diverse moral situations is commonplace. Moreover, psychological explanations are available as to why considered judgements vary within a society. As David Copp<sup>5</sup> points out:

"....one's sincere, complete and non-temporary confidence in a judgement can be explained by psychological mechanisms having to do with one's moral training and one's personality, and by sociological factors having to do with the culture of one's group, one's class background, and so on". (p.17)

In addition, the flexibility of individuals, in their willingness to revise considered judgements, in order to reach equilibrium, also varies considerably. This variation in flexibility can also be explained by psychological and sociological mechanisms.

These considerations seem to indicate that the set of considered judgements that a person would hold in wide reflective equilibrium depends critically on the psychology of a person and on sociological factors associated with the person's life. It is quite possible, for example, that due to traumatic childhood experiences, Jones comes to acquire sincere, complete and non-temporary confidence in the judgement that the subjugation of women is right. It is also entirely possible that theoretical considerations do not undermine his confidence in the above judgement, so that in trying to acquire an equilibrium, this judgement which we will suppose to be a considered judgement, can legitimately impugn theories as unjustified. Suppose furthermore, that Jones succeeds in reaching an equilibrium so that the judgement "Subjugation of women is right" is justified for him. Given what we know about the genesis of considered judgements, and the flexibility of persons in their willingness to revise judgements in order to attain equilibrium, I think we might be inclined to allow that Jones is justified in his belief that the subjugation of women is wrong. But the judgement itself may not be justified even if it is one that is held in wide reflective equilibrium. After all, Jones has come to acquire sincere and non-temporary confidence in this judgement via traumatic experiences. Notice what a conservative like Rawls or Daniels would be committed to saying if Jones' life history had been different. It is possible that if Jones' childhood had been a happy one, free of traumatic experiences, he would have acquired sincere and non-temporary confidence in the judgement that the subjugation



of women is wrong, and this judgement would have been held by him in wide reflective equilibrium. If so, the conservative would have to admit that the judgement "Subjugation of women is right" would not have been justified for him if only his childhood had been a happy one, and so on. But this gives us even more grounds for being inclined to say that even if  $j$  is a judgement that Jones would hold in wide reflective equilibrium, and even if he would be justified in holding  $j$ ,  $j$  itself might well be unjustified.

Of course, according to wide reflective equilibrium theory, in order for a judgement to be justified for a person  $A$ , it is not sufficient that the judgement merely be a considered judgement. The judgement must stand in a certain complex coherence relation  $R$  with background theories and principles. The fact that a considered judgement must stand in relation  $R$  with background theories and principles might lead us to believe that the question of whether a considered judgement is justified for  $A$  is not primarily a matter of  $A$ 's psychological idiosyncracies and of  $A$ 's sociological background. This fact might lead us to believe this because of the complexity of relation  $R$ : In order to attain equilibrium, relation  $R$  must be satisfied and the satisfaction of  $R$  demands theoretical argument and decision. A conservative might emphasize that one should not overlook the theoretical support that considered judgements have when they are judgements that would be held in wide reflective equilibrium. And on the basis of this, the conservative might be inclined to argue that what judgements would be justified for an agent in wide reflective equilibrium is not primarily a matter of the psychological and sociological constitution of the agent.

However, the fact that a considered judgement must stand in a certain complex coherence relation with background elements, or the fact that

considered judgements in wide reflective equilibrium have theoretical support, does little to vitiate the claim that what judgements are justified for a person, according to the R/D theory, is primarily a matter of the psychological and sociological constitution of the agent. Suppose  $j$  is a recalcitrant or an intractable considered judgement. That is to say, suppose  $j$  is a judgement that  $A$  is psychologically incapable of renouncing; it is a judgement that is deeply entrenched in  $A$ 's psychology. Then, according to the R/D theory, in order for this judgement to be justified for agent  $A$ ,  $j$  would have to be a judgement that  $A$  would hold in wide reflective equilibrium. Since  $j$  is a recalcitrant considered judgement, in order for  $A$  to attain equilibrium  $j$  must be accommodated within an equilibrium package. Furthermore, since  $j$  is an intractable considered judgement, let us assume that  $A$ 's confidence in  $j$  will not be tempered by any amount of theoretical consideration. In such a situation, regardless of whether  $j$  has theoretical support or whether  $j$  stands in relation  $R$  with background theories and principles,  $j$  remains a considered judgement. If  $j$  is accommodated within an equilibrium package,  $j$  will be a considered judgement with theoretical support and it will be justified. If  $j$  is not accommodated within an equilibrium package  $j$  will still remain a considered judgement but it will be unjustified. But given these considerations, according to the R/D theory, regardless of whether considered judgements have theoretical support, any judgement, in order to be justified, must first be a considered judgement. But since whether a judgement is a considered judgement ultimately depends on the psychology of the agent - it ultimately depends on whether the agent has complete, sincere and non-

temporary confidence in the judgement - it appears that whether a judgement will be justified for an agent in wide reflective equilibrium, ultimately depends on the psychological and sociological constitution of the agent. Theoretical support seems to be of secondary importance.

Finally, we indicate that although observation statements are used as standards against which non-moral theories are assessed, one does not need to rely ultimately on the psychology of the person who makes the observation in order for the observation statement itself to be justified. The set of observation statements that are justified ultimately depends not on the psychological idiosyncracies of observers, but on the way the world is. In the example we cited above (chapter 4), what ultimately explains why the physicist, under specific conditions, makes the observation "There goes a proton" is the existence of protons. Had the world been different, had there been no protons, the physicist would not have made this observation. He would not have made this observation since in the absence of protons, a vapour trail would, in all probability, not have been produced. However, assuming that protons exist, had the psychology of the physicist been different, had he, for example, had a happy childhood instead of an unhappy one, the physicist would still, under specific conditions, have made the relevant observation. This is because the psychological constitution of the physicist has no bearing on the existence or the non-existence of protons. Contrast this with the moral case: In the moral case, the considered judgements that a person would make under specified conditions depends critically on the psychology of the person. Vary the psychology of the person, and we vary the set of considered judgements that the person would make

under certain conditions. Unlike the making of observation reports, the making of considered judgements is not constrained, at least we lack good grounds for assuming that it is so constrained, by an external reality. Moreover, the absence of an "external reality" constraint on the set of considered judgements a person would make, "fits" well with the fact that there is considerable variation with respect to considered judgements that are made under similar conditions by different persons.

In summary, it appears that the justification of an observation statement does not ultimately depend on the psychological constitution of the observer, whereas according to the R/D theory, the justification of a moral judgement does ultimately, and primarily, depend upon the psychology of the person making the judgement. And it is this dependence that inclines one to the view that the R/D theory is best construed as a theory that specifies conditions under which a person is justified in his moral views.

REFERENCES

1. Copp, David, "Considered Judgements and Moral Justification: Conservatism In Moral Theory", forthcoming in David Copp and David Zimmerman, Morality, Reason and Truth, Dordrecht, Holland: Reidel, 1982.
2. Nagel, Thomas, "The Objectivity of Ethics", unpublished, 1980.
3. Brandt, Richard, A Theory of The Good and The Right, Oxford: Clarendon Press, 1979.
4. For more discussion on related examples, see Sturgeon, Nicholas L., "Brandt's Moral Empiricism", unpublished.
5. Copp, David, *ibid*, 1982, pp. 17.

## Chapter 6

### MORALITY AND SOCIAL INSTITUTIONS

In any complex society, numerous aspects of morality are embodied by social institutions. For example, aspects of morality such as those concerned with distributive justice, with abortion, with advertisement, with the sale of pornography, and with various environmental issues. If these institutions are to be just, or if they are to be subscribed to by all members of the society, then the moral principles underlying these institutions must themselves be justified. In the absence of any justification for these principles, the members of the society in question will lack moral reasons for subscribing to them. Of course, this is not to deny that they may have weighty prudential reasons for subscribing to them. In addition, if the moral principles underlying these social institutions lack justificatory support, then the institutions themselves, or the practices enjoined by these institutions, would be unjustified. These remarks are in accord with Rawls'<sup>1</sup> notion regarding the subject of justice: Not only is he interested in formulating principles of social justice but he is also interested in their institutionalization. Indeed he claims that:

"For us, the primary subject of justice is the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation. By major institutions, I understand the political constitution and the principal economic and social arrangements. Thus the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family are examples of major social institutions." (p.7)

In this chapter, I wish to explore some of the implications the R/D

theory has regarding the institutionalization of moral principles.

I will argue that the R/D theory allows for the possibility that two incompatible moralities be justified in relation to two subgroups within a particular society. I will take "society" as indicative of some very large group, for example, the group that constitutes Canadian or Japanese society. The possibility of multiply justified moralities within a society, in turn, I will argue, creates difficulties for the institutionalization of a justified set of moral principles.

### 1. Mutiply Justified Moralities

The R/D theory allows for the possibility that two or more equilibrium packages, or two or more justified moralities, will emerge in one society. Two equilibrium packages would arise in a society, for example, if the set of considered judgements, in equilibrium with principles and background theories, held by members of one subgroup of the society, were significantly different or incompatible with a second set of considered judgements, also in equilibrium with a different set of principles and theories, held by members of a second subgroup of the society. Such a situation is not improbable given what we know about the genesis of considered judgements: The considered judgements held by a person are a function of the person's psychological and sociological constitution. In addition, as we pointed out in chapter (5), the flexibility of individuals, in their willingness to revise considered judgements, in order to reach equilibrium, also varies considerably.<sup>2</sup> This variation in flexibility can also be explained by psychological and sociological mechanisms. Since the psychological and sociological factors germane to the genesis of considered judgements can and do vary from individual

to individual, it is reasonable to assume that the considered judgements held by different members of a society could vary both prior to and in equilibrium.

These considerations support the claim that the R/D theory allows for the possibility of multiply justified moralities, with respect to a particular society, or in the terminology that we will adopt, the R/D theory implicates the possibility of "intrasocietal" relativism.

Intrasocietal relativism would be implicated by a theory of moral justification if two conditions were true:

(i) The theory T justifies two different moralities, either simpliciter, or in relation to different persons or groups in the society. For example, T could be a theory according to which:

(a) Moral judgement or principle P is justified either simpliciter or in relation to A, where 'A' stands for some person or some group.

(b) Moral judgement or principle not-P is justified either simpliciter or in relation to B, where 'B' stands for some person or some group other than the person or group that 'A' stands for.

(c) Moral principles P and not-P are either both justified or both valid, or both justified or valid in relation to someone.

(ii) T does not specify a method of ethical reasoning that can be expected to show, when there is a conflict of principles, that one and only one principle can be justified in relation to all parties concerned, i.e., to everyone in the society, or that one and only one principle is correct or valid. Of course, not any method will do.

The kind of method referred to in (ii) would satisfy the following two conditions:



1. The non-arbitrary condition

The method should be non-arbitrary. It should not, for example, advocate the flipping of a coin in order to show that only one of the principles P or not-P is valid or correct, or that only one of these principles can be justified to all parties concerned.

2. The Justification condition

The method should be one such that if it is utilized, then one succeeds in specifying conditions under which a morality itself is justified, and not in specifying conditions under which a person or a group is justified in its moral views. Hence, condition (ii) is that T fails to specify a non-arbitrary method of objective justification.

Clearly the R/D theory implicates the possibility of intrasocietal relativism. We have seen that the theory, in connection with a particular society, allows for the possibility of multiply justified moralities, and the theory does not specify a method of ethical reasoning that would establish that only one of the moralities, would be valid or justified for everyone in the society.

Besides the R/D theory, various other meta-ethical theories implicate the possibility of intrasocietal relativism. Consider, for example, Richard Brandt's theory. Brandt argues that all fully rational persons would tend to support a moral code that has at least some elements - the so-called "Hobbesian Core" - that are commonly shared. Even granting this, it is plausible to assume that different rational agents in a society, apart from the Hobbesian morality, would tend to support different moral codes. Roughly, as indicated in chapter (5), according to Brandt's theory, given that an agent's desires are rational and given that an agent

has full relevant information, it would be rational for that agent to support a moral code. And since the moral code is supported by a fully rational agent, that code is justified for that agent. But if it is plausible to assume that different rational agents in a society would tend to support different moral codes apart from the Hobbesian morality, then Brandt's theory implicates the possibility of intrasocietal relativism. Jones, for example, might support a code that condemns abortion, whereas Smith might support a code that condones abortion. In this case, the judgement "Abortion is not wrong" would be justified for Smith, and the judgement "Abortion is wrong" would be justified for Jones. Moreover, Brandt's theory does not specify a non-arbitrary method of ethical reasoning that would establish that only one of these judgements is justified for both Jones and Smith.

## 2. Morality and Social Institutions

Ideally, if morality is to play a role "in governing the life or actions of a rational and informed person or society,"<sup>3</sup> the morality must be a justified one. Of course, even if a morality is not justified, it may play a role in guiding the life of a person. This would be the case, for example, if a person sincerely, but falsely, believed that a morality were justified for him. But in such a situation, if the person were to act in accordance with the morality in virtue of his believing that it is a justified one, his reasons for so acting would be based on a false belief. For such a person, morality would only inadvertently influence his life. Moreover, since a morality may be justified either simpliciter or in relation to a person or a group without being the group's morality in the sense that members of the group subscribe to it

or even know of it, moralities that are to figure in the lives of rational agents must ideally be both justified, and subscribed to by them. Let us distinguish between a morality that is both justified and subscribed to by some relevant group, and a morality that is a justified morality but that is not subscribed to by the relevant group in question.

We contended that the R/D theory implicates the possibility of intrasocietal relativism. It is clear that the theory would implicate the possibility of intrasocietal relativism even when the multiply justified moralities were subscribed to. I will now argue that any theory that implicates the possibility of intrasocietal relativism, where the multiply justified moralities are subscribed to, will encounter difficulties when it comes to the institutionalization of a set of justified moral principles. The argument can be summarized as follows:

- (1) Numerous aspects of morality are embodied in social institutions.
- (2) If these institutions are themselves to be justified, or if the practices enjoined by these institutions are themselves to be justified, then the moral principles underlying these institutions must themselves be justified.
- (3) In a situation where a theory of moral justification allows for the possibility of two would-be-subscribed-to moralities being justified, either simpliciter, or relative to subgroups of a society, there would be no appropriate method to decide between the institutionalization of incompatible sets of moral principles. The methods available would be inappropriate in that: (i) Either they would be arbitrary - they would violate the "non-arbitrary condition", or (ii) they would be methods, which if utilized, would succeed in specifying conditions under which

persons were justified in adhering to the practices of an institution, and not in specifying conditions under which the practices themselves were justified. In other words, the available methods would violate the "Justification condition".

(4) Given (3), any theory of moral justification that implicates the possibility of intrasocietal relativism, where the multiply justified moralities are subscribed to, will have difficulties when it comes to the institutionalization of moral principles. Since these difficulties arise specifically in virtue of the theory allowing for the possibility of intrasocietal relativism, the theory can have disturbing social implications.

I take it that premises (1) and (2) are not controversial. However Premises (3) and (4) require discussion.

Let us suppose that a theory of moral justification analogous to the R/D theory has succeeded in justifying two different moralities, in relation to different subgroups within a society. The two moralities need not be entirely incompatible. For example, they could agree on the prohibition of things like assault, rape, murder etc. but disagree significantly with regard to such issues as abortion, and distributive justice. Let us also suppose that a number of "social architects" have been commissioned the task of institutionalizing moral principles for the society as a whole. For example, it has been put to the social architects to institutionalize a complete and consistent policy of distributive justice, and a complete and consistent policy regarding the morality of abortion. Since the policies in question must be justified, the social architects begin by an examination of the prominent justified

theories of abortion and distributive justice. However, when they do this, they are confronted with the following problem: How are they to rationally decide between the institutionalization of the relevant principles of the two justified moralities, assuming that both these moralities have complete and consistent policies regarding distributive justice and abortion?

It is not open to them to decide by decree, or by flipping a coin since these methods would be arbitrary. R.M. Hare<sup>4</sup> (p.69) proposes an alternative. Hare seems to be suggesting that in a situation analogous to the one described above, a social architect should decide which morality to institutionalize, not by any arbitrary means, but by careful reflection on the consequences and the way of life represented by each of the moralities. Such a method to decide between the institutionalization of competing moralities would then reflect psychological facts about persons. This would be so since we know that the way of life favoured by an individual is a function of factors such as: whether the individual's upbringing was authoritative or liberal; whether the individual was brought up in a religious or in a non-religious environment; the social environment in which the individual spent his early childhood; the individual's education; and the early childhood experiences of the individual. Since these factors can vary quite radically from individual to individual, it is plausible to assume that persons with different psychologies would favour different ways of life and hence different moralities.

But if persons with different psychologies would favour different moralities, then different social architects with different psychologies, would favour different moralities. If two social architects were to

favour different moralities, it is unclear what non-arbitrary method could be propounded to decide between these competing moralities.

In addition, if we were to rely on such a procedure for choosing between the institutionalization of competing moralities, then we would ultimately be describing conditions under which a person were justified in his moral views and not conditions under which a morality itself were justified. This would be so since the choice between competing moralities would ultimately rest primarily on the psychological constitution of the chooser, and not on the content of the moralities in question: According to Hare's decision procedure, we would be claiming that given that a person is brought up in a certain environment, has a certain type of education, and has certain experiences, it is rational for him to prefer one lifestyle, or one morality, to another. But this decision procedure seems to leave it quite open as to whether the morality itself that a person would favour would be justified for him. Suppose, for example, as social architects, we decided that life would be altogether more pleasant if abortion were condoned, and we favoured the morality that differed from the alternate morality only in that it condoned, whereas the other condemned abortion. Then, as social architects, given our decision, we would perhaps be justified in our belief that abortion were permitted, but this would not imply that the moral judgement, "Abortion is permitted" were itself justified. Moreover, if the policy of permitting abortion were institutionalized, then as social architects it is possible that we would be justified in adhering to this policy or practice, but this would not imply that the practice itself were justified. To take another example, suppose, after careful reflection, as social architects

we decided that life would be better if benefits and burdens were to be distributed in accordance with utilitarian principles of distributive justice, rather than in accordance with the Rawlsian difference principle. Then as social architects we would perhaps be justified in our belief that utilitarian principles of distributive justice were just principles, but this would not imply anything about whether the utilitarian principles were themselves justified.

Is Hare's proposal the only available one to decide between the institutionalization of competing moralities, in a situation where a theory of moral justification has justified, either simpliciter or relative to subgroups, different moralities? Some theorists would claim that it is not. They would contend, for example, that in such a situation, one should adopt a "pluralistic" or a "liberalistic" policy with regard to institutionalizing moral principles. That is to say, in such a situation, institutions should be designed to accommodate a variety of different views about particular social policies. For example, Daniel Callahan<sup>5</sup> argues that:

"The strength of pluralistic societies lies in the personal freedom they afford individuals. One is free to choose among religious, philosophical, ideological, and political creeds; or one can create one's own highly personal, idiosyncratic moral code and view of the universe." (p.124)

In connection with abortion Callahan writes:

"It is reasonable and legitimate to say that a woman should be left free to make the decision in the light of her own personal values; that is I believe the best legal solution." (p.125)

Callahan then suggests a number of questions and considerations women ought "to think about as they try to work out their own views on abortion." (p.125) Questions such as: What philosophical assumptions are implicit in different uses of the word 'human'? What biological evidence should be considered? What philosophical theory of biological analysis is required? What are the social consequences of different kinds of analyses? (p.126)

In a society with multiply justified moralities, pluralists, in connection with the issue of institutionalizing moral principles, could argue in the following way:

"The question of rationally deciding between the institutionalization of competing moral principles need not arise. It need not arise since one can design social institutions which embody principles general enough to accommodate a variety of views regarding such issues as abortion, advertising, distributive justice, etc. For example, in connection with abortion, one could institutionalize highly permissive laws which specified that the final decision is to be left to the woman."

Is pluralism a credible alternative to Hare's proposal when it comes to the institutionalization of competing moralities?

First we should note that it is not a genuine alternative to Hare's proposal since it does not offer a rational procedure to decide between the institutionalization of competing principles. According to our interpretation of pluralism, no such procedure is required.

Second, in accordance with our construal of pluralism, it is unclear whether it would be of any help in institutionalizing moral principles if the principles in question were not general enough. For example, how would pluralism be of any help in rationally adjudicating between the institutionalization of the justified principles "Abortion is right"



and "Abortion is not right"? The pluralist might reply that in such cases, persons should be left to decide on their own, and the principles that are institutionalized should reflect this position.

But this reply will not do. Firstly, it seems to be highly question-begging. Secondly, with respect to certain moral issues, it is not even prima facie plausible. For example, in the event that a theory of moral justification justified utilitarian principles of distributive justice relative to one subgroup, and Rawlsian principles relative to a second subgroup, it would be implausible to permit individuals to choose, perhaps via the medium of highly permissive laws, whatever system of distributive justice they were partial to.

Finally, this reply is inadequate since, if one were to adopt what it recommends, then one would violate the "Justification condition": Consider a society with highly permissive abortion laws, where the final decision is left to the woman. Even if a woman, in deciding whether to obtain an abortion, were to reflect carefully on all the relevant considerations outlined by Callahan, her decision would largely be influenced by her conception of the good life. After careful reflection, perhaps she would be justified in her belief that abortion is right, but this would not justify the belief that abortion is right. Like Hare's proposal, the pluralist's suggestions as to how one should proceed with regard to institutionalizing justified moral principles in a society with multiply justified moralities, violates the "Justification condition".

The difficulties that would arise in institutionalizing moral principles if a theory of moral justification were to justify two or more would-be-subscribed-to moralities, either simpliciter or in relation

to subgroups, for a society, would arise specifically in virtue of the theory justifying more than one morality for the society. The difficulties would not arise, for example, if a theory of moral justification justified a single morality that was subscribed to, either simpliciter or in relation to persons, for the society as a whole. In this case, it would be the moral principles of the unique justified morality that would ideally be embodied in the various social institutions.

If the above argument is cogent, then it could be extended to cases involving inter societal relativism. That is to say, if we had a set of social institutions relevant to all societies and specific to no particular society, then with respect to this set of institutions, difficulties, similar to the ones discussed above, would arise, regarding the institutionalization of competing moralities, with a theory of moral justification that implicated the possibility of inter societal relativism. Universal defense systems and universal health-care systems would be examples of the types of institutions that we have in mind.

To summarize, we have presented an argument to the effect that a theory of moral justification that can justify two or more moralities that are subscribed to by different groups, for a society, and that does not specify a rational procedure to establish that only one of the moralities can be justified for all members of the society, will encounter difficulties in institutionalizing moral principles. The argument will not appear compelling to a person who cares little about rational decision making: Such a person, when confronted with the issue of institutionalizing incompatible justified moral principles, may eschew rational decision

making. He may opt for coin-flipping. But if rationality is an underlying constraint, and if the moral principles at the foundation of our social and political institutions are to be both rational and justified, then one cannot ignore the troublesome social consequences that are entailed by a theory of moral justification that implicates the possibility of intrasocietal relativism.

### 3. The Uni-Moral Constraint

The entailment of these untenable consequences, by a theory of moral justification that allows for the possibility of intrasocietal relativism, suggests a general constraint that it would be reasonable to accept in enabling one to decide between competing theories that purport to justify a moral judgement or a morality. Let us call this constraint the "Uni-moral constraint". Our defense for this constraint simply rests on the ideal of the rational society. As mentioned above, if rationality is an underlying constraint, and if the moral principles at the foundation of our social and political structures are to be rational and justified, then the Uni-moral constraint recommends itself as a regulatory constraint.

We will conclude this chapter by developing the Uni-moral constraint. Roughly, according to this constraint, given two (or more) theories of the justification of a morality, all other things being equal, it would be reasonable to prefer the theory that does not implicate the possibility of intrasocietal relativism.

When would "all other things be equal"? Good meta-ethical theories must satisfy a number of constraints. Amongst the not-so-very controversial constraints, we can list the following:

- (i) The theory must be consistent and non-circular.
- (ii) It must satisfy conditions of simplicity - it must not multiply entities unnecessarily.
- (iii) The theory must explain how it is that moral language can be used to recommend.
- (iv) The theory must explain the phenomenon of reason giving in moral discourse.

I do not pretend that this list is complete. Some theorists, for example, argue that meta-ethical theories must not have normative implications. But suppose we have circumscribed a reasonable set of conditions that an ethical theory must satisfy,  $a_1, a_2, \dots, a_n$ , excluding the Uni-moral constraint. Then all other things would be equal between two theories of moral justification if they satisfied the same conditions from the set.

To illustrate and to elaborate some of these ideas, consider theories of moral justification  $T_1$  and  $T_2$ . Let  $T_1$  be a theory according to which:

- (i) Moral judgement or principle P is justified either simpliciter or in relation to A, where "A" stands for some person or some group.
- (ii) Moral judgement or principle not-P is justified either simpliciter or in relation to B, where "B" stands for some person or some group other than the person or group that "A" stands for.
- (iii) Moral principles P and not-P are either both justified or both valid, or both justified or valid in relation to someone.

We will assume that the principles P and not-P conflict in some sense of "conflict".

Let  $T_2$  be a theory according to which:

(i) Moral judgement or principle P is provisionally justified either simpliciter, or in relation to A, where "A" stands for some person or some group.

(ii) Moral judgement or principle not-P is provisionally justified either simpliciter, or in relation to B, where "B" stands for a person or group.

Let us also assume that  $T_2$ , unlike  $T_1$ , specifies a method of ethical reasoning that can be expected to show, when there is a provisional conflict of principles, that one and only one principle can be justified to all parties concerned, or that one and only one principle is correct or valid. The method must satisfy the "Non-arbitrary" and the "Justification" conditions.

Suppose  $T_1$  and  $T_2$  both satisfied the same conditions that good meta-ethical theories must satisfy. Then the Uni-moral constraint would have it that it would be reasonable to prefer  $T_2$  to  $T_1$ .

We can now formulate a final version of the Uni-moral constraint.

The Uni-moral constraint:

Given two theories of moral justification  $T_1$  and  $T_2$ , where:

(i)  $T_1$  and  $T_2$  both satisfy the same conditions from the set  $a_1, a_2, \dots, a_n$ , where  $a_1$  to  $a_n$  are conditions that good meta-ethical theories must satisfy, and

(ii)  $T_2$  differs from  $T_1$  in that  $T_2$  satisfies condition M, whereas  $T_1$  does not, it would be reasonable to prefer  $T_2$  to  $T_1$ .

Condition M: If the theory allows for the possibility that it provisionally justifies different, incompatible moralities for different subgroups in a society, the theory specifies a method that will show that one and only

one of the moralities is justified for everyone in the society. The method satisfies the "Non-arbitrary" and the "Justification" conditions. Recall that the Non-arbitrary condition states that the method should be non-arbitrary, and the Justification condition requires that the method be one such that if it is utilized, then one succeeds in specifying conditions under which a morality itself is justified, and not in specifying conditions under which a person or a group is justified in its moral views.

REFERENCES

1. Rawls, John, A Theory of Justice, Cambridge: Harvard University Press, 1971.
2. See: Copp, David, "Considered Judgements and Moral Justification: Conservatism In Moral Theory", unpublished, 1982, p.17.
3. Ibid, p.2.
4. Hare, R.M., The Language of Morals, New York: Oxford University Press, 1964, p.69.
5. Callahan, Daniel, "Abortion: A Matter of Conscience" in Raziel Abelson and Marie-Louise Friquegnon (eds.), Ethics For Modern Life, New York: St. Martin's Press, 1975.

## CONCLUSION

In chapter (1), we contended that conservatism, as an approach to developing and assessing moral theory, commands a prevalent position on the current philosophical scene. The work of Rawls and Daniels represents the most ambitious and sustained attempt to developing a conservative theory of moral justification. Although it is impossible to remain unimpressed by the ingenuity and sophistication of the theory, conservative wide reflective equilibrium has its pitfalls.

In this thesis, we have tried to indicate the major inadequacies of the theory. We generated three independent criticisms of the theory. First, we questioned the conservatism of the theory: We argued that neither Rawls nor Daniels seems to have a rationale for using considered judgements as standards against which ethical theory is to be assessed. In addition, the inability of the theory to circumvent the "no credibility" objection suggests that conservative wide reflective equilibrium theory cannot be viewed as a comprehensive or general theory of moral justification: The theory must assume ab initio that our considered judgements, the very standards against which ethical theory is to be assessed, are themselves justified.

Secondly, we argued that the R/D theory is not a theory germane to the justification of a morality. This argument turned on the distinction between a person's being justified in holding a morality and a morality itself being justified. We presented considerations to sustain the view that, at best, the R/D theory specifies conditions under which persons are justified in their moral beliefs.

Finally, we argued that the R/D theory has a potential for disturbing



social implications - it is a theory susceptible to difficulties when it comes to institutionalizing moral principles. These difficulties arise in virtue of the theory implicating the possibility of what we called "intrasocietal relativism".

The cogency of these criticisms will dictate whether we are to reassess the tenability of the R/D conservative wide reflective equilibrium theory.

BIBLIOGRAPHY

1. BONJOUR, Laurence, "The Coherence Theory of Empirical Knowledge", Philosophical Studies, 30, 1976.
2. BRANDT, Richard B., Ethical Theory, Englewood Cliffs, N.J.: Prentice Hall, 1959.
3. BRANDT, Richard B., A Theory of The Good and The Right, Oxford: Clarendon Press, 1979.
4. CALLAHAN, Daniel, "Abortion: A Matter of Conscience" in Raziel Abelson and Marie-Louise Fricquegnon (eds.), Ethics For Modern Life, New York: St. Martin's Press, 1975.
5. COPP, David, "Considered Judgements and Moral Justification: Conservatism in Moral Theory", forthcoming in David Copp and David Zimmerman, Morality, Reason and Truth, Dordrecht, Holland: Reidel, 1982.
6. DANIELS, Norman, "Wide Reflective Equilibrium and Theory Acceptance in Ethics", The Journal of Philosophy, 76, 1979.
7. DANIELS, Norman, "Reflective Equilibrium and Archimedean Points", Canadian Journal of Philosophy, 10, 1980.
8. DWORKIN, Donald, Taking Rights Seriously, Cambridge: Harvard University Press, 1977.
9. FIRTH, Roderick, "Ethical Absolutism and the Ideal Observer", Philosophy and Phenomenological Research, 12, 1952.
10. HARE, R.M. The Language of Morals, New York: Oxford University, 1964.
11. HARMAN, Gilbert, The Nature of Morality, New York: Oxford University Press, 1977.
12. HEMPEL, C.G., Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, New York: The Free Press, 1965.
13. NAGEL, Thomas, "The Objectivity of Ethics", unpublished, 1980.
14. PAMPAS, George S. (editor), Justification and Knowledge, Dordrecht: Reidel, 1979.
15. PAMPAS, George S. and SWAIN, Marshall (editors), Essays on Knowledge and Justification, Ithaca: Cornell University Press, 1978.

16. POLLOCK, John L., "A Plethora of Epistemological Theories" in PAMPAS, George S. (ed.), Justification and Knowledge, Dordrecht: Reidel.
17. RAWLS, John, A Theory of Justice, Cambridge: Harvard University Press, 1971.
18. RESCHER, Nicholas, "Foundationalism, Coherentism, And The Idea of Cognitive Systematization", The Journal of Philosophy, 71, 1974.